数据防泄露(DLP)技术指南

中国信息协会
信息安全专业委员会

天空卫士
SkyGuard

目录

一、数据防泄露（DLP）与数据安全治理 7

1、数据安全必须先做数据安全治理 7

2、数据防泄露（DLP）是数据安全治理的重要目标导向 8

二、数据防泄露（DLP）的定义 9

三、数据防泄露（DLP）的应用 9

1、数据防泄露（DLP）为实现数据分类分级保护提供技术支撑 9

2、数据防泄露（DLP）在数据生命周期提供安全防护 10

a）识别存储数据中敏感信息的位置 11

b）掌握流转中数据的数据流向及流转方式 11

3、对使用中的数据进行权限细分，并进行相应的控制 12

a）终端数据使用 12

b）应用数据使用 12

四、数据防泄露（DLP）的核心技术 13

1、方案部署 13

a）分布式部署架构（分层管理） 13

b）策略分级部署 13

c）证据文件外置部署 13

d）管理服务器高可用 14

e）数据库高可用 14

2、策略定制 14

a）内容检测的组合检测 14

b）检测对象的锁定与过滤 14

c）策略响应动作 15

d）策略触发通知 15

3、数据识别 15

a）关键字识别 16

b）字典权重识别 16

c）正则表达式识别 16

d）文件属性识别 17

e）图像识别 17

f）标签识别 17

g）机器学习识别 17
h）指纹识别 18
i）其他识别 18
4、管理与控制 21
a）统一管理控制台 21
b）策略多模块分发 21
c）预置策略 21
d）角色管理 21
e）多权分立 22
f）策略审批部署 22
g）接口集成 22
h）用户识别 22
5、分析与报告 22
a）事件处理 23
b）事件日志 23
c）事件工作流 23
d）格式化显示 23
e）事件关联合并 23
f）日志查询功能 24
g）基于用户的报告 24
h）数据分类查询 24
五、数据防泄露（DLP）的应用场景 24
1、存储数据防泄露（发现 DLP） 25
a）风险分析 25
b）关键功能 27
2、网络数据防泄露（网络 DLP） 28
a）风险分析 28
b）关键功能 29
3、邮件数据防泄露（邮件 DLP） 30
a）风险分析 30
b）关键功能 31
4、终端数据防泄露（终端 DLP） 32

a）关键功能 34
5、应用数据防泄露（应用 DLP） 37
a）风险分析 37
b）关键功能 39
6、移动数据防泄露（移动 DLP） 40
a）风险分析 40
b）关键功能 42
六、数据防泄露（DLP）的扩展 44
1、人工智能与行为分析加入数据防泄露（DLP）体系 44
2、数据安全治理的自动化平台 45
七、术语与定义 46

一、 数据防泄露（Data Loss Prevention，简称 DLP）与数据安全治理

1、数据安全必须先做数据安全治理

随着数字产业化和产业数字化的快速推进，数据已经成为数字化转型时代的核心竞争力。随着数据成为资产，成为基础设施，数据的安全受到前所未有的重视和保护。数据安全治理融合了数据安全技术和数据安全管理，是以“数据使用安全”为目标的技术体系。通过对数据安全治理的必要性以及其发展现状进行分析，提出了一种以数据安全标识为基础的数据安全治理体系框架和技术架构。《中华人民共和国国民经济和社会发展第十四个五年规划和 2035 年远景目标纲要》明确提出：数据是推动数字化发展的关键要素，必须加快数字产业化和产业数字化，推进数字化发展，推动数字经济和实体经济深度融合。

为更好地发挥数据的基础资源作用和创新引擎作用，要做好数据资源的开发、利用和保护：一方面要使数据连起来、跑起来、用起来；另一方面要强化数据的安全保障。数据安全治理是“围绕数据安全使用”的愿景，以“让数据使用更安全”为目的的安全体系构建方法论，覆盖了敏感信息管理、安全防护、合规三大目标而构建的技术体系。它不仅是一套多种数据安全工具协同组合的产品级解决方案，而且是从管理制度到工具支撑，从决策层到技术层，自上而下贯穿整个组织架构的完整体系链条。

数据的安全和使用是相互矛盾而又需要调和的两个方面，如何在确保数据高效使用的前提下，又保证数据的安全管理，成为一个值得关注的问题。

a） 数据规模暴涨，战略价值提升

根据《大数据白皮书（2020 年）》统计，2020 年全球共产生数据量达到 47ZB，预计到 2035 年这一数据量将达到 2142ZB，全球数据量逐年迎来爆发式的规模增长，数据已成为一种能够影响国家战略决策的战略资源，谁掌握了更多、更有价值的数据，谁就掌握了未来的主动权。

b） 数据泄露频繁，数据安全需要治理

根据 ForgeRock《消费者身份信息违规报告》2020 年公布的数据，网络犯罪分子在 2019 年暴露了超过 50 亿条数据记录，给美国造成了超过 1.2 万亿美元的损失。随着数据泄露问题的日趋严重，对数据安全治理的需求越来越迫切。

2、数据防泄露（DLP）是数据安全治理的重要目标导向

在企业数字化转型的过程中，业务的核心驱动是资产化的数据。这些资产化的数据，是企业高度依赖甚至决定企业业务战略可实现性的关键元素。这些数据资产，是独立的黑客组织抑或生态的竞争对手，都期望获取的核心价值资产。

数据安全治理，就是为了在企业数字化转型期间，精确的通过科学化、流程化、自动化、弹性化，高度灵活来适配企业的业务流程，把数据资产的管理与基于生命周期的安全通过技术能力，与业务逻辑进行深度的融合。从而实现企业核心数据资产的安全性保护，并最大程度的在企业战略目标与风险容忍度之间，达到最佳的动态平衡，让企业在数字化活动中获取最大业务收益，同时实现合规遵从。

对数据资产的泄露防护，就是数据安全治理流程体系所输出的能力检验标准之一，也是业务数据安全建设的关键技术支撑能力。所以做数据安全治理必须首先建设数据防泄露体系。

二、数据防泄露（DLP）的定义

数据防泄露（DLP）指的是使用先进的内容分析技术，在统一的管理控制台内对静止的、流转的、使用的敏感数据进行保护的系统，是当前支撑数据资产保护的重要技术之一。

DLP 产品分为两个类型：企业级 DLP 和集成 DLP。企业级 DLP 技术专注于防止敏感数据丢失并实现多数据通道防护的全面覆盖，集成式 DLP 解决方案则是在其他非 DLP 设备中提供有限的 DLP 功能。

许多安全厂商都认可数据泄露防护的技术重要性，但不少厂商采用集成 DLP 解决方案，导致了许多机构和管理者被误导，认为数据泄露防护解决方案的防护范围和能力不够充分。所以本文将制定一套标准规范体系，帮助组织机构构建完善的数据防泄露解决方案。

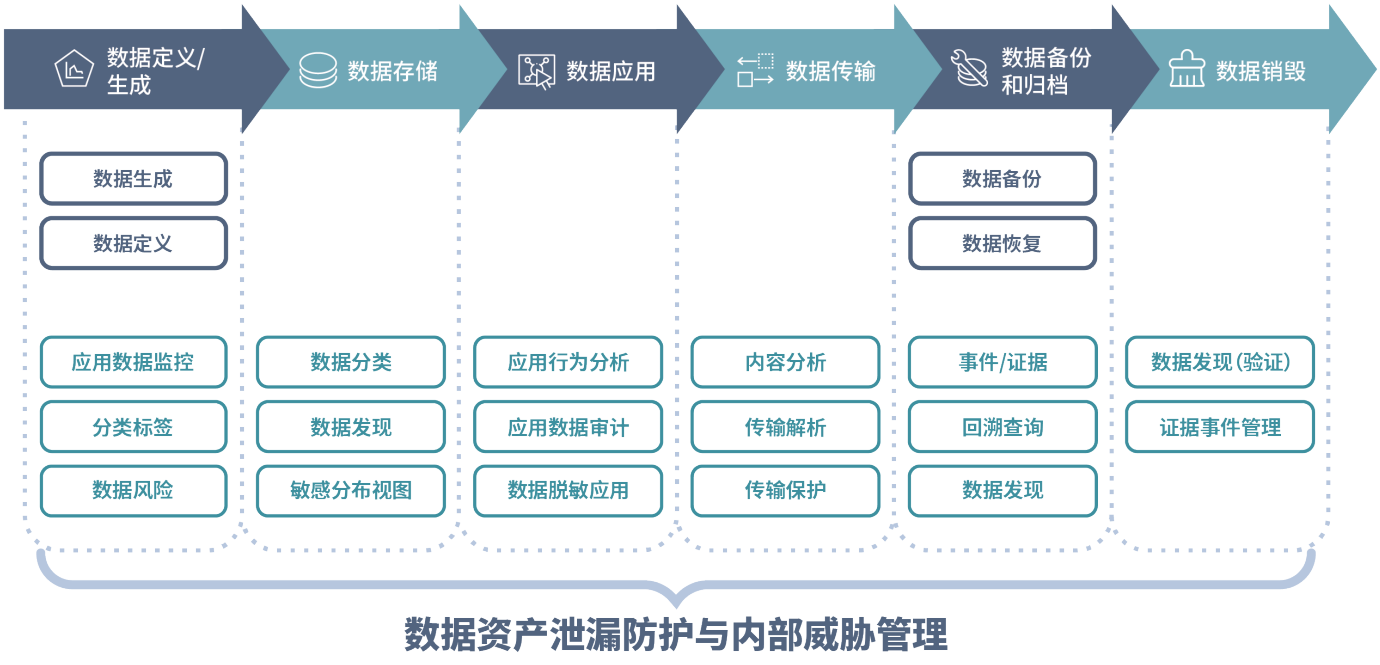
三、数据防泄露（DLP）的应用

1、数据防泄露（DLP）为实现数据分类分级保护提供技术支撑

我国《中华人民共和国数据安全法》第二十一条：“各地区、各部门应当按照数据分类分级保护制度，确定本地区、本部门以及相关行业、领域的重要数据具体目录，对列入目录的数据进行重点保护。”所以，数据分类分级是数据安全的基础。通过将数据分组到合理的“数据类别”中，同时，对数据的重要程度进行标记，将对整个企业的数据保护和自动化控制提供良好的基础。

如果缺乏数据分类分级，就会对数据的风险缺乏了解，数据可能会受到不正确的管理和对待。DLP 不仅仅是解决数据泄露风险的技术方法，还是控制和减轻业务风险的手段。通过对数据内容的检测分析，为数据的分级分类提供有效的技术支撑。DLP 应被引入到业务风险管理流程中，从根本去解决未经授权或不安全的数据转移，从而确保数据防护技术与业务需求保持一致。

2、数据防泄露（DLP）在数据生命周期提供安全防护



数据防泄露贯穿于数据生命周期的全过程。从数据的生成开始，到数据的存储、应用、传输、备份归档到数据的销毁，每一个步骤都有相应的数据防泄露技术作为支撑。

在数据生命周期的保护中，数据防泄露最重要的功能是能对数据按照分类分级进行识别，并针对不同级别和类型的数据在不同的阶段采取不同的保护措施。

a）识别存储数据中敏感信息的位置

当今是数字化转型时代，数据存储在各种应用程序和设备中，存储位置多样、数据量庞大，DLP 帮助安全人员了解重要数据的存储位置。

b）掌握流转中数据的流向及流转方式

DLP 帮助安全人员了解数据的流向以及流转方式，包括第三方数据访问、业务数据交换和终端数据流转等。数据流转方式可以分为“网络流转”、“应用流转”、“第三方数据共享”和“终端流转”四种。

① 网络数据流转

在所有流转方式中，网络数据流转是最主要的数据泄露途径。这里主要应关注利用常见的网络协议与外界进行通讯的行为，包括：

- ✓ 通过 HTTP / HTTPS 协议上传数据；
- ✓ 利用 SMTP 邮件传输协议外发内部邮件；
- ✓ 使用 FTP 文件传输协议上传文件；
- ✓ 使用 SMB 协议对外上传文件等。

② 应用数据流转

应用系统包含了集中存放的数据，是最核心的数据资产所在。多数应用系统以 HTTP/HTTPS 方式对外提供服务，数据使用者涵盖了内部用户和外部客户，涉及大量的非结构化文档的交互，数据在应用流转的过程中，极其可能因为敏感数据暴露面扩大而产生数据泄露问题，存在敏感数据不正当扩散的风险。

③ 第三方数据分享

除了应用访问和内部流转外，根据《网络安全法》《数据安全法》和《个人信息保护法》等国家和相关行业。监管机构的法规要求，安全人员还应掌握第三方访问敏感数据的交换信息，确定可以访问敏感数据的第三方名单，采取相应的控制措施来确保第三方访问的数据安全，防止敏感数据被过度过量获取。

④ 终端数据流转

终端是最终用户日常操作数据的重要平台。内部人员可以通过终端的各种通道，如: 移动存储、网页、邮件、应用、光盘刻录、打印、文件共享等对内、外进行数据流转的动作，以实现办公和业务的相关流程。

3、对使用中的数据进行权限细分，并进行相应的控制

a）终端数据使用

内部的敏感数据在未进行完整保护的情形下可能被下载到内部人员的 PC 和笔记本电脑，用于合法办公或业务使用当中。有权访问端点的任何员工或外部人员都可以将所有或部分这类数据下载到 DVD、U 盘驱动器甚至 iPad 上，并针对敏感数据进行复制、粘贴、打印、截屏等相关的操作。

b）应用数据使用

关键性敏感数据产生于业务和应用系统，存在人员权限滥用、 违规操作、越权访问和下载等风险，需要通过人员权限细分、控制、审计和考核来降低核心数据操作风险。根据企业安全合规要求，敏感数据应该实行最小化授权。不合理的数据不应该出现在企业应用中，敏感数据不应该被越权用户下载。

因此围绕数据的不同形态，可以有效降低风险的全面数据安全解决方案必须能够准确地发现存储在文件服务器、电子邮件系统、应用程序存储、关系型数据库或其他数据存储库中的敏感数据。并合理利用 DLP 的内容识别能力，围绕数据流转及使用等情况，对于敏感数据的异常操作行为进行控制，对敏感数据的不正当扩散进行审计和阻止。

四、数据防泄露（DLP）的核心技术

1、方案部署

a）分布式部署架构（分层管理）

DLP 系统应支持在多个监控位置部署，这些监控节点应该能够将所有 DLP 事件发送回中央管理服务器以便形成工作流、报告、调查和归档，应支持分层部署，支持在不同区域的管理服务器上运行不同区域的策略。

b）策略分级部署

DLP 系统应支持对下属机构设置独立的管理员，对所管理的区域对像放开策略设置、事件查看、报告查看等功能；管理员的功能模块不可以超出上级管理员；为了满足统一策略集中管理的需求，总管理员可以向下属机构进行策略推送，子管理员可以对总管理员推送的策略进行查看，但无法进行编辑和删除的动作。

c）证据文件外置部署

DLP 系统应能够对数据泄露事件提供证据文件外部保存能力，可采用外置文件存储（NFS/SMB）的方式进行集中存储。

d）管理服务器高可用

DLP 系统应支持使用两台管理服务器进行主备模式部署，如果主服务器关闭或服务不可用，备用服务器将自动接替行使管理功能。

e）数据库高可用

DLP 系统应支持数据库服务集群部署，通过虚拟 IP 技术使数据库节点保持高可用状态，如果主数据库活动节点关闭或不可用时，备用节点将接管活动角色，入库日志、事件、证据、配置等保持同步且不丢失。

2、策略定制

a）内容检测的组合检测

由于 DLP 系统中存在众多内容识别分类器，每个分类器均可以独立配置作为内容识别的手段。但在有些情况下需要通过采用多个检测策略规则进行搭配，支持基于“与”、“或”、“非”这三种判断的各种组合关系，减少误报的几率。

b）检测对象的锁定与过滤

通常，DLP 的检测对象可以涵盖两类对象，分别是检测对象和检测通道。检测对象一般指发送数据来源 / 目的、邮件地址、组织架构等，而检测通道则泛指网络通道和终端通道这两种通道类型。利用 DLP 的检测对象过滤功能，应更加精准地锁定检测的范围和目标，使得检测结果更加准确。

✓ **基于对象的检测**

在实际使用 DLP 系统时，由于部分员工的特殊性要求或正常业务所需，可以使用 IP、域名、邮件地址和组织 架构对检测对象进行合理的限制（通俗上称为：策略的黑白名单）。

✓ **基于通道的检测**

在 DLP 定义中，DLP 系统应支持对各个模块进行统一策略的分发，这样就需要 DLP 支持管理员应可以对不同通道下的不同协议或存储通道进行不同的操作，从而达到细化策略的效果。

c） 策略响应动作

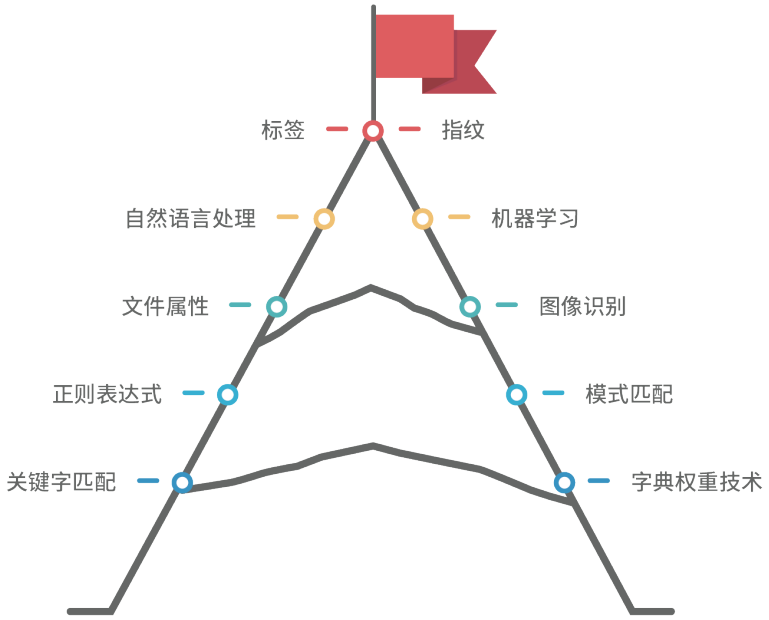
策略的响应动作是指当流量或执行的动作命中了预设的检测策略后，针对不同的通道在网络侧或终端侧执行的动作，通常情况下的动作执行包括但不限于放行、阻止、隔离、确认、个人密钥加密等。

d） 策略触发通知

在运维工作中，当最终用户的动作命中了检测策略，管理员应可以在不登陆 DLP 系统的情况下及时获知触发策略的事件行为。DLP 系统可以通过执行发送邮件通知，提醒特定人员等相关事件的发生。通知的内容应该支持定制和常见的变量引用，如：企业特定的 Logo、公司名称、邮件主题、正文内容等。

3、数据识别

数据识别技术是 DLP 解决方案中使用各种技术分析深层内容的能力。当前全球主流 DLP 产品所支持的数据识别技术根据其匹配敏感信息的精准程度，至下而上依次为：**关键字识别、字典权重识别、正则表达式识别、文件属性识别、图像内容识别、自然语言分析处理、标签识别、机器学习识别** 和 **指纹识别** 共计 9 种。检验 DLP 产品是否具有完备及可用的数据识别技术是检验 DLP 产品最核心的功能指标。



a）关键字识别

关键字识别分类器是目前 DLP 数据识别器种类中最基本的功能。采用关键字其特点为技术实现简单，配置容易。DLP 系统应支持使用关键字匹配分类器对敏感数据进行检测。

b）字典权重识别

字典权重分类器一般是为了方便引用相同类型的关键字信息字典。使用字典权重分类器除了方便引用外一般还需要考虑到特定的词语在整个检测信息中的实际权重。DLP 系统应支持使用字典权重识别分类器对敏感数据进行检测。

c）正则表达式识别

正则表达式是对字符串操作的一种逻辑公式，是用事先定义的一些特定字符及这些特定字符的组合，组成 “规则字符串”，从而实现对字符串的一种过滤逻辑。利用正则表达式识别器的数据检测方式可以完成自定义产生或常见的具有一定规则的字符串的检测。

d）文件属性识别

文件属性识别分类器并不会对检测的文件进行内容检查，该分流器主要关心的是检测文件自身所携带的属性，包括：文件类型、文件大小、文件名称、附件数量等。

文件类型识别通常可以用来检测外发的文件是否为压缩、加密或特定的文件类型，一旦检测出此类文件类型，企业则可以采取对应的措施加以处理。

文件大小识别通常还用来判断外发的文件的实际大小，并结合其他识别器以提高检测的精确度。

e）图像识别

通过光学字符识别技术，在 DLP 识别分类器中可以提取包含在图片中的文字，并将提取出的文字用于后续的内容识别。

f）标签识别

通过在格式化文件，如：微软 Office、PDF、视频、音频、图片等文件中的属性编辑，这些分类分级标签可以对在 DLP 的各个传输通道中的数据进行识别并根据标签含义采取相应的保护手段。

g）机器学习识别

机器学习识别分类器是指将一批相似的同类文件提交给 DLP 系统进行学习，DLP 系统采用机器学习算法，根据上下文和自然语义提炼出这些文件的共同点，并将这些共同点按照算法生成类别特征，用于对后续不同通道中发现的被检测文件进行判别。

h）指纹识别

指纹识别分类器目前主要分成非结构化指纹（文件指纹）和结构化指纹（数据库指纹和 CSV 文件指纹）两种分类器。文件指纹是指使用指纹识别分类器扫描样本文件内容，根据相关技术做成指纹信息，对被检测文件进行相似度匹配；而数据库指纹是指：使用指纹识别分类器扫描数据库表里的每个单元格信息，从而实现对被检测内容进行数据库指纹匹配。通常情况下，指纹识别具备较高的精度，可以准确判断出被检测文件是否包含敏感内容。

✓ **文件指纹采集**

指纹样本的采集应该定期更新，所以 DLP 产品应满足样本文件自动上传至服务器并自动进行新样本的指纹生成 的要求。DLP 指纹采集方式至少应该支持压缩包上传和远程目录自动定期读取两种方式。

✓ **文件特征离线式采集**

文件特征的生成需要 DLP 管理员对文件目标有访问权限。对于特殊行业或特殊情况下 DLP 管理员对原始目标不具备访问权限，应提供辅助工具，由对文件目标有权限的人员进行指纹或机器学习的特征生成，然后通过定时导入功能自动并定期将产生的特征文件导入服务器内预防二次泄密。

i）其他识别

✓ **模式匹配识别**

模式匹配识别是指数据在与外界交互过程中，企业关注检测对象的哪种行为。在防止内部数据泄露的场景下，安全人员一般关心内部用户提交的数据（如 HTTP Post 方法）。而在保护服务器数据的模式下，则需要 DLP 产品支持 HTTP Response 数据的内容检测。利用模式识别技术可以形成对敏感数据外发检测和敏感数据异产品支持 HTTP Response 数据的内容检测。利用模式识别技术可以形成对敏感数据外发检测和敏感数据异常下载两种场景的检测闭环。

✓ **内容简繁体自动转换检测**

简繁转换是在检测匹配的过程中，检测内容既检测简体中文内容也同时检测对应的繁体中文内容。这样就要求 DLP 系统将检测内容提取文字信息后，能够自动完成简体文字和繁体文字的互相转换的过程。

✓ **压缩内容的检测**

在实际业务场景中，大量存在将文档进行压缩后进行传输外发的行为。因此，DLP 系统需要具备正确识别多重压缩文档里的文件内容和类型的能力。

✓ **文件嵌套的检测**

为了防止最终用户采用在正常的 Word 文档中嵌套含有敏感数据文档的方式逃避 DLP 检测，DLP 系统需要支持对文件嵌套的内容检测。

✓ **变形文件的检测**

为了防止最终用户采用通过修改文档后缀的方式逃避 DLP 的类型和内容检测，DLP 系统需要支持对变形文件的类型和内容检测。

✓ **加密文件的检测**

为了预防内容识别为核心的 DLP 系统无法识别加密内容导致的数据泄密，DLP 系统需要支持对加密文件类型的识别，防止敏感数据通过加密文件外发的行为。

✓ **自定义文件的检测**

特殊行业尤其是制造行业使用的应用程序所产生的文件内容通常不能被 DLP 系统所识别。需要 DLP 系统具备能将同类文件通过自身系统或附加工具进行自我分析后找到相同文件特征信息的能力。通过使用分析后的文件特征，DLP 系统就能具备识别外发该类文件的能力。

✓ **压缩分片的检测**

为了识别使用压缩软件将敏感文档进行压缩分片外发规避检测的行为，DLP 系统需要具备正确识别“使用压缩软件进行分片压缩的文件类型”的能力。

✓ **零星式泄露的检测**

组织在实际业务场景中需要对敏感数据（如：身份证信息和手机号等）以少量多次的泄露方式进行识别。DLP 产品应提供一定时间内，基于多次泄露的状态统计。如 60 分钟中内，外发的内容达到一定的数量则立即阻止，及时避免此类泄露。

✓ **聚类技术**

采用非监督学习聚类算法，将大量混杂的文件数据按照文件内容的相似度自动聚类，并通过机器学习生成类别特征信息。协助数据管理者进行数据分类分级，对聚类后的文件可以基于字典、正则表达式、指纹、智能学习等技术，辅助数据管理者有效的制定 DLP 防护策略。

✓ **自定义数据模板**

通过可编程控制技术，基于特定的文件类型和内容格式，对被检测文件进行分析和处理，精准的判别文件的类型和分类等级，可以有效辅助数据管理者在复杂的环境下精准的对检测文件进行识别。

4、管理与控制

a）统一管理控制台

全套 DLP 解决方案的一个独特而关键的功能是完善的中央管理控制能力，应可以管理覆盖包括“发现、网络、邮件、终端、应用、移动”等所有 DLP 技术架构下的安全组件，从而实现对移动数据、静止数据和使用中数据的覆盖范围、创建和管理策略、报告和事件工作流进行相应的管控。

b）策略多模块分发

策略多模块分发是指最终用户不需要为各 DLP 模块设置不同的检测策略，通过控制台即可集中为 DLP 产品各

模块下发统一的检测策略 , 也可以单独为某个 DLP 产品模块或模块组合下发独立策略。

c）预置策略

预置模板是指为了方便 DLP 用户更加便捷的使用数据防泄露产品，DLP 系统在内部提前给使用者定制好检测内容的检测模板。预置策略模板应该是开箱即用并能对外发的数据进行有效且精确的识别。

d）角色管理

DLP 系统应该允许基于角色的内部管理来进行内部管理任务以及监视和执行。在内部可以将用户分配到管理和策略组以分离职责，为监视和执行提供灵活的支持，使之能投入到不同的策略管理工作中。

e）多权分立

系统应支持多权分立方式登陆管理平台，包括：系统管理员（负责系统管理），安全管理员（负责审批管理），审计管理员（负责审计管理），事件管理员（负责事件审计）

f）策略审批部署

在实现分权管理后，DLP 系统应支持使用特定角色对发布的检测策略进行有效性稽核及审批生效的工作，满足了大型机构对 DLP 管理权限分离的需求，降低了因错误 DLP 策略对生产或办公业务产生影响的可能性。

g）接口集成

DLP 系统应支持在组织内对接特定的集成产品，包括并不局限于 SIEM、SOC、甚至是安全自动化运维平台等，具备细颗粒度的事件外发能力。同时 DLP 系统应支持与运维系统进行对接，通过 API 对策略中的来源和目标进行增加、删除和修改的动作。

h）用户识别

事件管理需要合理利用组织内的用户信息，将所有违规者与用户身份数据关联起来，DLP 系统应可以支持基于用户的认证（Active Directory）登录，从而使 DHCP 租约信息与企业登陆用户联系起来，将身份信息的上下文关联添加到每个事件告警中，避免将策略违规与正常用户联系在一起。

5、分析与报告

DLP 系统应具备提供实时告警及对受保护数据实时分析的能力，及时通知组织内安全人员有关泄露或违规事件的信息，并根据需要采取手动操作。这对于涉及到核心数据资产外泄或严重的违规事件的及时处理特别有用。其次，分析和报告功能可帮助 DLP 管理员密切关注数据的整体安全性以及解决方案的效能，应能提供相应的合规报告，以确保组织满足相应的合规要求。

a）事件处理

事件处理队列，是指分配给 DLP 管理员进行事件处理程序但事件仍处在处理阶段的事件摘要。每一个事件可以对任何字段进行排序或过滤，包括通道，违反策略，用户，事件状态和处理流程。

b）事件日志

DLP 系统应具备记录单个事件的详细信息能力，包括但不限于事件类型、发生时间、上报时间、发送者、接收者、违反策略、告警级别等内容。

c）事件工作流

DLP 系统应具备多种工作流程，所有工作流记录都必须包含：发送者、接收者、传输方式、所涉及内容的类型、内容的安全等级以及实际内容本身。这将为安全团队提供调整策略，纠正数据滥用和跟踪潜在高风险用户所需的相关信息。

d）格式化显示

DLP 系统应支持对使用 Webmail 外发敏感数据时，能够在事件详情内智能格式化展现邮件发送人、邮件接送者、邮件抄送者、邮件密送者、邮件主题等相关信息。

e）事件关联合并

DLP 系统应支持对短期内同一来源及同一目标和同一外泄方式的相似事件进行合并展示，降低管理员处理事件的复杂性，减少问题处理时间。

f）日志查询功能

DLP 系统应具备完善的日志查询功能，支持按不同参数进行查询、过滤和排序报告。可以通过相应的过滤器，对所有的日志进行精确查询和报告，例如：用户组、策略组、策略规则、严重性、用户名等。

g）基于用户的报告

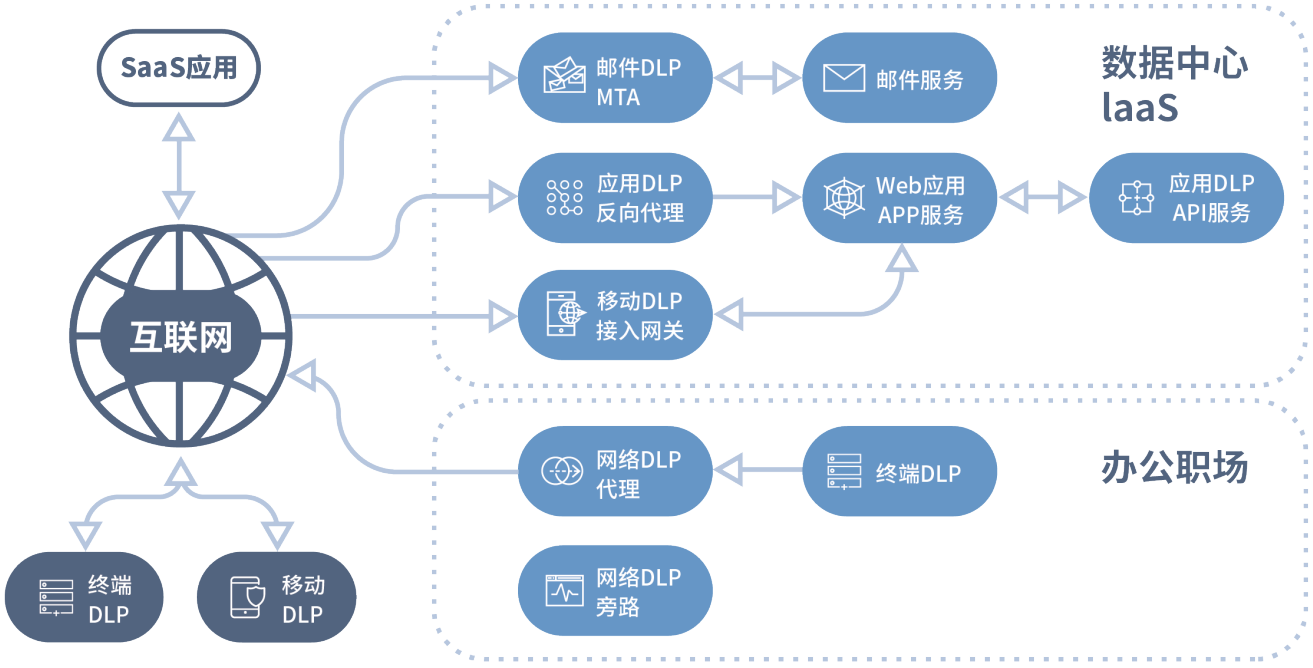
DLP 系统应可以针对个人风险值、个人风险排名、个人事件数据统计等生成以人为中心的报告。

h）数据分类查询

DLP 系统应提供基于“数据分类”并且以数据分类的形式来呈现数据安全报告。

五、数据防泄露（DLP）的应用场景

DLP 解决方案既能保护敏感数据，又能深入了解组织内数据的使用情况。DLP 帮助安全人员更好地理解数据，并提高其分类和管理内容的能力。如下图所示：全套 DLP 解决方案需要提供对整个组织的网络、邮件、数据库、移动应用、端点、内部业务应用的全面覆盖。本节的其余部分将重点讨论专用的 DLP 解决方案能力和常见的 DLP 架构。



在数字化转型的形势下，数据无处不在，因此 DLP 的部署应该是覆盖完整的企业 IT 架构，包括职场网络、数据中心 /IaaS、移动办公终端、SaaS 应用等。在统一的数据安全策略下，这些位置的 DLP 部署应当能保持联动，在统一的数据安全策略下保护组织的数据资产。

1、存储数据防泄露（发现 DLP）

a）风险分析

① 存储数据违规风险

数据存储是指组织机构在提供产品和服务、开展经营管理等活动中，将数据进行持久化存储的过程，包括但不限于采用云存储服务、网络存储设备等载体存储数据。而且随着业务的不断发展，存储位置也在不断的增加，如测试网、生产网、办公网等，都有可能存储敏感数据，如果不能掌握和了解数据的存储分布情况，就做不到对数据的精准保护。所以，如何要对存储中的数据进行保护，前提就是需要对数据存储的分布情况进行梳理，得到内部的数据分布视图。

数据存储过程，可能存在敏感数据泄露、数据篡改、数据丢失、数据不可用等安全风险。许多行业数据保护法规要求限制终端或人员对敏感信息的访问，并且仅允许在限定的时间内保留这些数据。最重要的是，许多数据主体现在有权要求删除其数据或可以选择撤销对数据处理的同意。如果组织不知道敏感数据存储在公司内的哪些位置，他们就有可能违反数据保护法规，并因不合规而面临巨额罚款。此类风险行为包括：

✓ 需要加密保存的数据未加密保存；

✓ 需要脱敏的数据未脱敏存储；

✓ 生产数据被放到测试网络；

✓ 高密级数据存储在低密级区域；

✓ 员工终端违规保存敏感数据；

✓ ……

② 数据删除 / 销毁检测

数据删除 / 销毁是指在停止业务服务、数据使用以及存储空间释放再分配等场景下，对存储位置中的数据进行清理的过程。机构需要采取安全控措施确保数据被安全有效的销毁。而数据存储对于此类要求可以进行定期的检查，防止应该清理的数据没有被及时销毁的风险。

通常，组织一般采用发现 DLP 来解决上述风险。发现 DLP 主要用于识别、定位和分类敏感或受监管的静止数据，以确定需要保护的受影响资源以及潜在的数据泄露风险，是合规的重要技术组成部分。数据发现使安全团队可以识别这些数据从而对其进行保护，并确保其机密性、完整性和可用性。

采用基于上下文内容感知的数据发现带来的优势包括：

✓ 增强了组织对拥有的数据以及存储位置的掌控能力，从而加强敏感数据访问和传输方式的理解过程；

✓ 利用发现 DLP 功能对存储的数据基于上下文的自动数据分类技术实现重要数据发现；

✓ 在发现敏感数据后执行触发补救事件以保护敏感数据；

✓ 协助完成组织内完整的敏感数据可见性，降低审计成本；

✓ 了解所有存储敏感信息的位置，在敏感数据被移动到未经批准的位置时持续进行监控。

b）关键功能

✓ DLP 系统应支持通过管理其他 DLP 模块的同一台服务器来管理内容查询功能，使用与其他 DLP 策略一致的静态数据扫描策略，从而保持策略、工作流程和事件处理的一致性；

✓ DLP 系统应支持发现扫描存储在数据库中敏感数据，包括 Oracle、SQLServer、MySQL、Postgres、DB2 等数据库，以及支持使用 ODBC 协议连接各类数据库进行扫描发现；

✓ DLP 系统应支持各类邮件应用系统，包括 Exchange、Outlook PST、Lotus Domino 进行敏感数据发现扫描；

✓ DLP 系统应支持对网络存储数据的敏感内容发现检测；

✓ DLP 系统的数据发现策略应支持计划定期扫描、增量与全量扫描、实时控制扫描，并可针对扫描的结果执行隔离或复制的补救动作。

2、 网络数据防泄露（网络 DLP）

a）风险分析

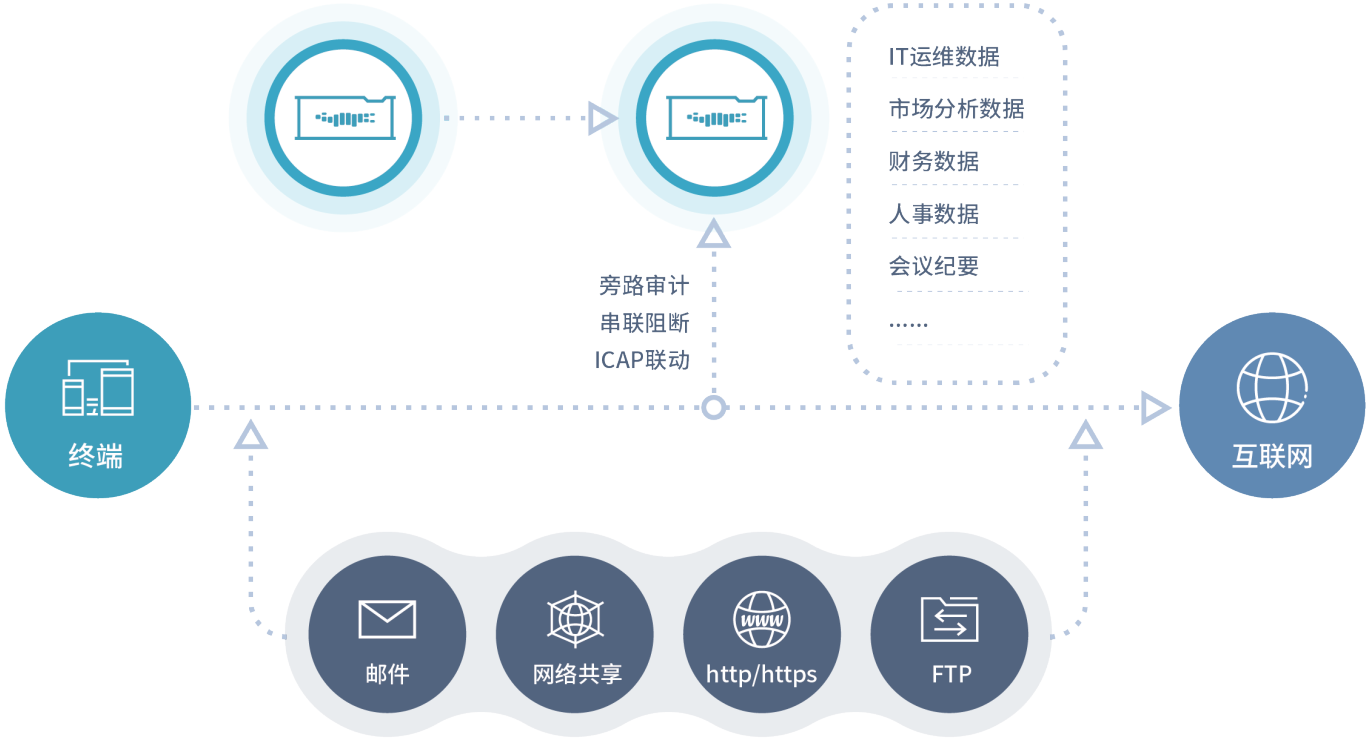
网络承载了数据的传输工作，维系着业务逻辑和办公人员的办公行为，保障网络敏感数据传输的安全尤为重要。由于各组织的职责不同，同一类数据在不同组织的安全级别存在较大差异，需要针对不同的组织制定不同的数据防护策略，包括：

✓ **跨域数据暴露风险**

跨域数据的共享面临的风险是高等级安全域向低等级安全域开放其无权访问的数据，在数据共享时未加以处理和区分，从而导致敏感数据泄露到低安全域内。随着企业数据共享的范围增加，数据安全的责任主体就随着共享范围的增加而扩大，新的数据拥有者往往容易忽略数据安全合规的要求，导致敏感数据泄露。

✓ **高威胁互联网数据行为风险**

开放的互联网工作方式导致安全性复杂化并增加风险。企业面临可能接触到敏感数据的内部人员包括企业的运维人员、操作人员、数据使用者可能会无意或蓄意利用网络环境将核心敏感数据泄露至安全网之外，令信息安全管理部门防不胜防、无从查证。



网络 DLP 通常以专用硬件设备或软件形式部署在网络边界或逻辑网络信任区至另一个网络区域的过渡点处。网络 DLP 既可以通过旁路监听的方式，也可以串联或代理的部署方式，同时支持多个网络 DLP 设备进行集群化部署。

b）关键功能

- ✓ 支持多种部署模式，要求支持旁路部署、代理模式、串行 (inline) 部署、ICAP 模式以适应各种场景；
- ✓ 可支持从旁路切换串联，产品无需重新安装；
- ✓ 支持在 IPv6 架构下部署和使用；
- ✓ 支持 HTTP/HTTPS、SMTP、FTP、IMAP、POP3 协议的敏感信息识别与监控；
- ✓ 支持对上传给外部网站的数据进行敏感内容检测和阻断；
- ✓ 支持对从企业内部服务器或云应用 Portal 中下载到用户电脑中的内容进行敏感内容检测，并对禁止下载的内容进行报警或阻断；
- ✓ 支持对特定 URL 类别进行内容检查和保护；
- ✓ 支持 SSL 流量解密；
- ✓ 支持 OCR 检测；
- ✓ 支持串联的自动 Bypass；
- ✓ 支持 ICAP 模式（代理捕获流量，将其发送到 DLP 产品进行分析，并在出现违规时终止通信。这意味着假设已经使用了 ICAP 兼容的网关，不需要在网络流量之间添加另一个硬件，而且 DLP 供应商可以避免构建用于内联分析的专用网络硬件）。

3、 邮件数据防泄露（邮件 DLP）

a）风险分析

DLP 产品的下一个主要组件是邮件 DLP。电子邮件系统是当前我国企事业单位最重要的信息沟通手段之一，其最重要功能就是对外进行通讯。电子邮件系统可能涉及的数据安全风险主要存在以下两种：

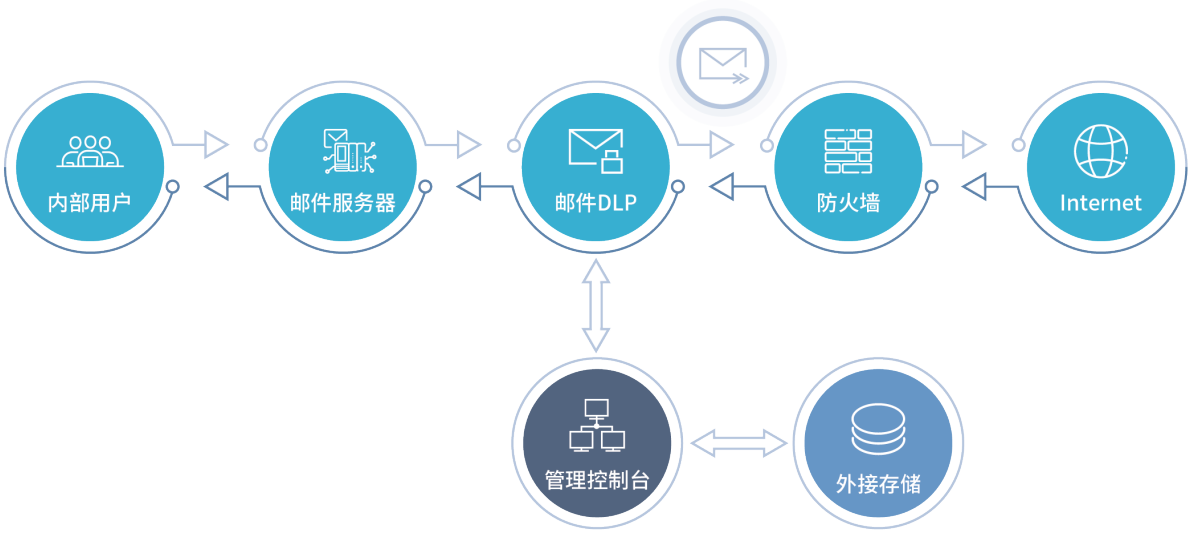
① **保密性风险**

保密性风险是指未经授权泄露保密信息，散播组织内部专有信息或业务信息，从而导致遗失知识产权；包括采用组织内部邮件对外进行敏感数据传输，同时也包括通过内部邮箱和移动设备邮件程序进行未经授权的敏感数据的邮件传递行为。

② **对外策略风险**

违反电子邮件正常使用策略的企事业单位人员，包括利用电子邮件系统发表个人言论带来的纠纷，以及内部用户出现未能遵守国家或企业的内审部门的监管要求使用邮箱，带来的违反保密规定、骚扰员工、违背合同责任、损害商誉和违反法律法规等问题。

由于邮件系统的特殊性，相关的邮件可以作为对应的追责证据进行举证，要求组织使用邮件 DLP 对于外发邮件的内容进行检查、审计和控制，这样就可以在出现安全事件时对于邮件内容进行追溯，快速的进行定位。



b）关键功能

“保密性风险”和“对外策略风险“常见的保护措施包括：

✓ **邮件链路加密**

邮件 DLP 应支持在特定的邮件通讯启用邮件的链路加密功能，对于外发邮件的传输过程及流量进行加密传输，防止监听和截取的风险行为出现。

✓ **邮件内容一致性检测**

邮件 DLP 应支持对于外发邮件内容，包括收发件人、主题、正文、附件等进行合规性检查。

✓ **邮件全量审计**

邮件 DLP 应支持对于所有外发的邮件进行审计留存，方便后续定向追溯。

✓ **邮件外发审批**

邮件 DLP 应支持对于一致性检测失败的邮件进行审批复核，在审批通过后再次外发。

✓ **邮件违规隔离**

邮件 DLP 应支持对于一致性检测失败且复核失败的邮件，视作违规邮件，对其进行主动隔离处置，防止数据外泄并留档核验。

✓ **邮件自动加密**

邮件 DLP 应支持对于包含敏感数据但仍有正当外发需求的邮件自动进行加密，防止数据外泄。该加密区别于链路加密，指的是对于邮件本身的加密。

✓ **邮件反向追溯**

邮件 DLP 应支持基于特定的邮件内容对全量审计的邮件基于邮件正文、附件、收发件人等内容进行快速检索和检查，从而缩小检测范围。

✓ **邮件定向追溯**

邮件 DLP 应支持在获得泄密邮件后，通过附件标记完成该文件的责任确认，锁定最后外发人和相关邮件信息，实现法律取证的可行性。

4、终端数据防泄露（终端 DLP）

a）风险分析

由于数据资产主要分布在业务系统和办公终端电脑中。这些终端电脑都是办公人员花费大部分工作时间并访问敏感信息的设备，也是数据丢失或被盗的主要来源。对终端主机而言，导致数据泄露的途径很多，主要可能存在以下风险：

✓ **终端数据采集风险**

数据采集是指某些组织机构在提供产品和服务、开展经营管理等活动中，直接或间接从个人信息主体、企业客户及外部数据供应方采集数据的过程。当数据产生后，由于业务分析、开发测试等需要，需要从生产系统批量下载和导出数据，这就是终端数据泄露的主要源头。在这个过程中存在人员权限滥用、违规操作、越权访问和下载等风险，需要通过人员权限细分、控制、审计和考核来降低核心数据操作风险。

✓ **终端数据存储风险**

数据存储是指结构或组织在提供产品和服务、开展经营管理等活动中，将数据进行持久化存储的过程。数据存储过程，可能存在敏感数据泄露、数据篡改、数据丢失、数据不可用等安全风险。另外，许多数据主体有权要求删除其数据或可以选择撤销对数据处理的同意。如果不知道敏感数据的存储位置，就有可能违反数据保护法规，并因不合规而面临巨额罚款。

✓ **终端数据使用风险**

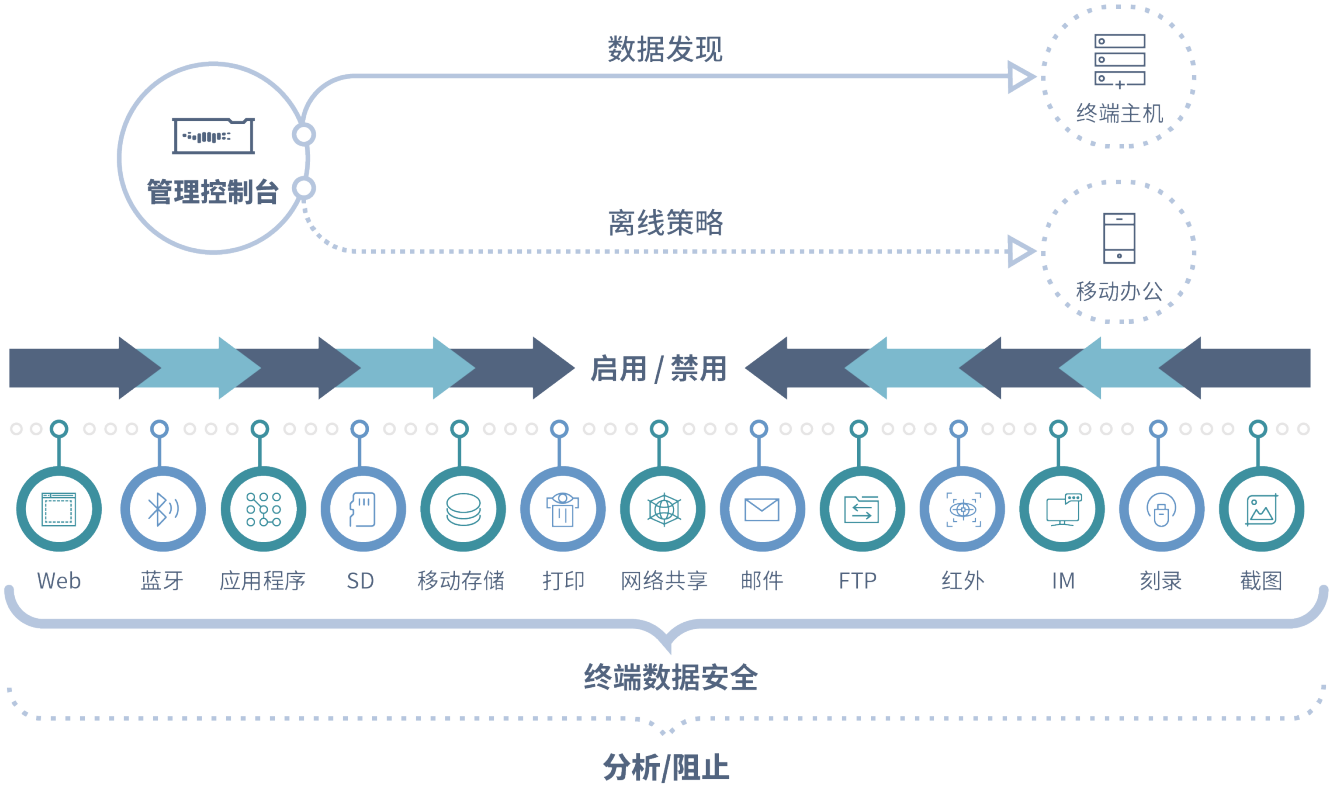
数据使用是指单位在提供产品和服务、开展经营管理等活动中，对数据进行访问、导出、以及分析、脱敏、加工、清洗等处理的过程。其中在终端上的数据使用的表现形式是指用户当前正在与之交互的数据，包括涉及敏感数据的屏幕捕获、敏感数据内容的复制 / 粘贴、敏感数据内容打印和传真操作等有意或无意通过非网络通信渠道传输敏感数据的尝试。

✓ **终端数据传输风险**

数据传输是指单位将数据从一个实体传输到另一个实体的过程，数据传输涉及与其所关联的全通信网络架构，可分为内部数据传输和内外部机构间的数据传输两种形式，当敏感数据流转到终端后，主要的泄露途径有互联网传输、外设传输和应用传输三种。

可以通过安装在用户的终端的 DLP 软件，防止企业的核心数据资产以违反安全策略规定的形式流出。终端 DLP 应全面兼容 Windows，Mac OS 操作系统，对终端上的文件共享、邮件、Web、应用程序等传输的数据

进行监控，在数据进行操作之前对其进行管控，如：打开、另存、复制、打印、拷贝、刻录、在线输入信息、在线传输文件（如：QQ、微信、邮件）等，并根据安全策略产生相关的动作（如阻止、审计、提示、水印、添加标签等），同时生成预警日志和审计日志，能够最大化的加强对关键数据的管控。



b）关键功能

✓ 终端支持类型

终端 DLP 应支持 Windows 全系列客户端，支持 MacOS 和国产化 Linux 进行部署，对上述操作系统均能实现基于内容的检测审计和阻断敏感数据外发。

✓ 终端协议数据检测

终端 DLP 应支持对使用 SMTP、HTTP/HTTPS、FTP、SMB、POP3、IMAP、ActiveSync 等协议进行传输的数据内容进行有效的识别，并根据策略对应的动作执行阻断、审计放行、弹窗提示等相应操作。

✓ 终端应用数据检测

终端 DLP 应支持对使用应用（包括云应用）客户端传输的数据内容进行有效的识别，并根据对应策略执行阻断、审计放行、弹窗提示等相应操作。

✓ 终端 IM 数据检测

终端 DLP 应支持对使用终端即时聊天工具，如 QQ、微信、飞书、钉钉、Skype 等聊天的文字与传送的文件内容进行检测与拦截阻断，终端 DLP 应可以获得接收和发送人员信息，支持自动截屏保留证据。

✓ 终端邮件数据检测

终端 DLP 应支持对使用 Foxmail、Outlook 等邮件客户端发送的邮件、通过命令行发送的邮件进行内容分析，并支持放行、阻断、确认动作。

✓ 终端离线策略

终端 DLP 应满足感知能力，能够判断终端所处的场所位置加载对应的检测控制策略，并对离开组织管控范围的受控终端执行更加严格的离线控制策略进行数据保护。

✓ 终端存储策略检测

终端 DLP 应支持对通过网络共享拷贝文件至本地存储的内容分析，并支持放行、阻断、确认动作。

✓ 终端浏览器插件的数据检测

终端 DLP 应支持对通过浏览器下载文件的行文进行 DLP 内容分析和策略匹配，防止敏感数据落到用户的计算机上。

✓ 终端加密数据检测

在终端文件加解密或者终端透明加解密系统的配合下，终端 DLP 应支持对加密的文件进行内容提取，并进行 DLP 内容分析和执行 DLP 策略。

✓ 终端打印通道

终端 DLP 应具备对打印的文件进行敏感内容分析检测及后续的阻挡控制能力，预防终端用户利用打印方式将敏感数据带离安全环境。

✓ 终端显性水印能力

终端 DLP 应支持显性水印的能力，包括屏幕水印能力、指定 URL 的浏览器水印、基于内容的应用程序水印以及打印水印等。

✓ 终端隐形水印能力

终端 DLP 应支持屏幕隐性水印，可对完全不可见水印的用户截屏进行反向解析获得用户信息。同时也支持打印暗水印，可以比较隐蔽的方式在打印件上随机分布附加图形，实现对员工的打印文件进行追溯。

✓ 终端手动标签

终端 DLP 应支持终端对文件进行手动标记标签的方式，为文件附加数据分类分级的标签，从而建立企业数据特征为后期检测使用。

✓ 终端应用程序标签

终端 DLP 应支持终端主机在本地使用 Office、WPS 等办公软件编辑文档后进行保存时，自动弹窗提示数据编辑者为该文件附加数据分类分级的标签，建立企业数据特征为后期检测使用。

✓ 终端自动扫描标签

终端 DLP 应支持通过数据发现的能力，在终端上将匹配特定敏感内容的文件自动打标签，为该文件附加数据分类分级的标签，建立企业数据特征为后期检测使用。

✓ 终端下载标签

终端 DLP 应支持对特定 URL 或共享目录所下载的文件自动打标签，为该文件附加数据分类分级的标签，建立企业数据特征为后期检测使用。

✓ 通讯加密

终端 DLP 通过网络上传事件到 DLP 管理服务器的过程中，应保证传输数据包的安全性。即使数据包被中途截获也需要确保内容无法查看。

✓ 配置加密

为了预防恶意用户在获知检测内容和匹配规则后进行规避，应确保收到检测策略后终端 DLP 保存在本地配置文件的保密性。

5、 应用数据防泄露（应用 DLP）

内部数据资产主要分布在业务系统和办公终端电脑中，其中应用系统包含了大部分集中存放的数据。这些业务系统既能提供给外部客户使用，也可以提供给内部用户使用；这些应用多数以 HTTP/HTTPS 方式对外提供服务，大量的数据会被录入到业务系统中，同时，业务系统也在输出各种数据以提供服务。这些数据往往是最核心的数据资产所在，一旦发生泄密，将导致所存放的客户的隐私数据或者组织自身的知识产权等相关内容的泄露，造成巨大的经济损失和组织的名誉损失，更加严重的情况还将受到刑法的处置。因此对应用系统的数据安全防护已经到了一个刻不容缓的时刻。

a） 风险分析

在传统的应用的数据安全中，大多数解决方案都集中在针对外部的防护方面。但基本上这类安全防护的手段无法解决以下风险：

✓ 应用系统的数据上传风险

应用系统本身要接收大量的用户提交数据，这些数据可能是用户的交易信息，查询信息等，也可以是用户交流信息内容。在传统安全范畴内，基本上对用户的上传内容很难进行分析和审计，这些违规内容的上传，会给企业带来巨大的影响和伤害。因此，需要在企业业务系统的内容上传过程中，对用户提交的内容进行内容分析和检测，发现其中的不正当内容，并对其进行阻止和拦截。

✓ 应用系统的数据读取 / 下载风险

尽管采用了很多的对外防护安全手段，但实际上大多数的应用系统为内部使用的，在内部用户有权限的情况下，或者当黑客绕过了所有的外部攻击手段后，就可以顺利的将内部的数据进行下载。

大多数的应用系统对服务器返回的数据并没有进行审查，同时对获取的数据内容、传输数量等没有进行控制和限制，因此，需要通过必要的手段在网站下行位置对服务器的返回内容进行分析和检测，发现其中可疑的操作及时进行审计和阻止。

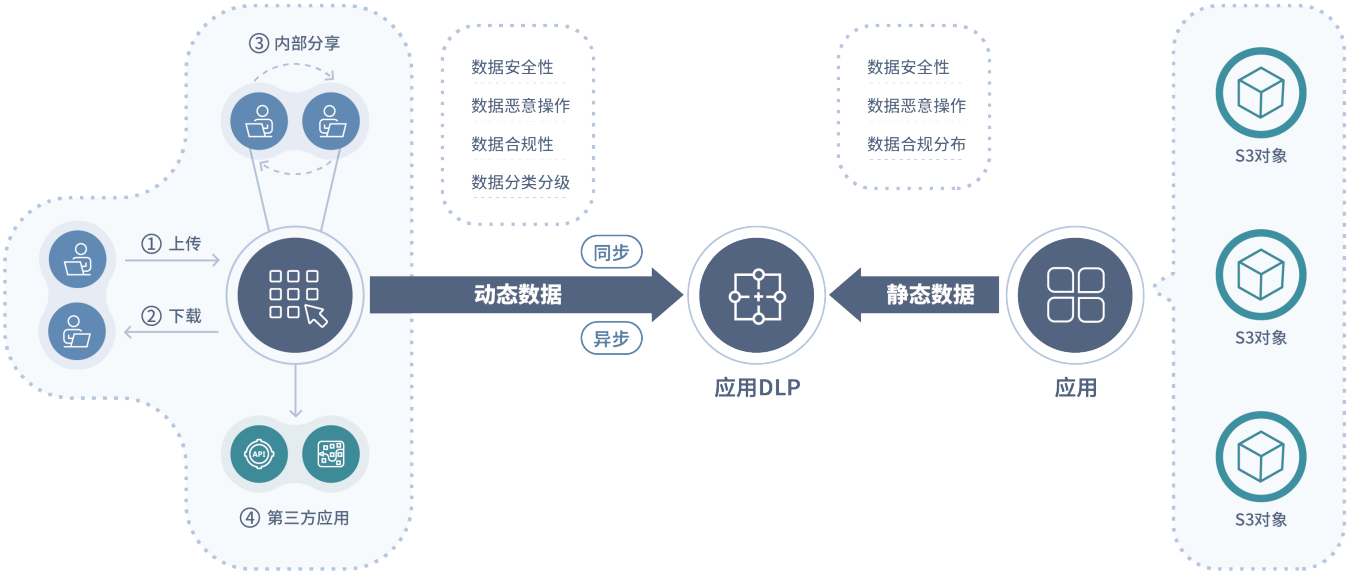
✓ 应用系统的存储数据风险

在很多的业务系统中，存在有大量的非结构化文档的交互，比如网盘系统，用户可以自由的上传和下载文档，也可以非常容易的将这些文档进行共享。而接收共享的人无需系统的用户名和密码就可以下载这些文档，这样，在一定的程度上，存在敏感内容不正当扩散的风险。因此，需要有必要的手段，增强文档管理系统的内容识别能力，在用户上传、下载、共享的时候对内容进行识别，对其中的敏感数据的不正当扩散进行审计和阻止。

✓ 应用系统与第三方系统数据交互风险

实际上，组织拥有的供应商越多，数据泄露的可能性就越高，只要有一个第三方系统出现问题就能造成数据泄露。因此，组织需要保护敏感数据，就需要考虑与第三方合作的风险。安全管理者需要了解整个组织的第三方生态系统，以及有哪些敏感数据与供应商进行交互。

因此，组织应该通过部署接口形式的“应用 DLP”设备，对内部业务应用自动进行内部调用数据、调用方、接收方、时间等信息，通过 RESTful 方式发送给“应用 DLP”设备的接口，由该设备对该数据进行深度上下文内容和安全风险识别分析。



b）关键功能

✓ **应用送检支持**

应用 DLP 应支持采用同步或异步模式，使得第三方应用可以通过 REST API 的方式将所要扫描的内容或文件发送给应用 DLP 并获取同步匹配策略的结果（阻断 / 放行），同时也应支持对上传至应用 DLP 的文件进行病毒检测，确保文件安全性。

✓ **应用存储送检支持**

应用 DLP 应支持第三方应用将对象存储方式、存储的数据发送给 UCWI 进行内容分析。可支持检测多种云存储对象并展示存储对象的区域、路径等属性信息。

✓ **文件脱敏支持**

应用 DLP 应支持对包含敏感内容的文件进行脱敏，自动去除敏感信息并返回脱敏文件给应用，便于安全操作。

✓ **信息获取支持**

应用 DLP 应提供包括事件列表、事件详情、证据文件、事件报告等相关命中策略信息，以 REST API 的方式提供给第三方应用进行调用读取。

6、移动数据防泄露（移动 DLP）

a）风险分析

随着信息化进程的推进，大量员工开始进行移动化办公，使用手机、pad 等移动设备进行邮件的收发、业务数据的处理等工作。然而，由于移动办公融合了移动通信、智能终端、信息技术，与传统电子办公相比有很大区别，移动办公的安全问题不仅包含了传统电子办公系统的安全问题，还包含了移动化引入的许多新的安全问题与隐患。这些新的风险主要包括以下几方面：

✓ **移动终端违规接入用户业务内网**

由于移动办公涉及公共运营网络，若缺乏有效的移动终端安全控制措施，则存在来自非法终端或恶意用户对办公信息系统进行非授权或异常访问的风险。

✓ 移动终端无线信道潜在的窃听风险

若在移动办公的过程中，通过无线信道传输的信息未加密，则存在严重的窃听风险。

✓ 移动终端的离散化特性加大了数据的监管难度

由于移动终端位于用户身边，地理位置相对分散，因此加大了移动终端及其数据访问管理、跟踪审计的难度。若缺乏必要管控手段，则可能出现监管死角，造成办公数据泄露。

✓ 移动终端易丢失、更换频繁

由于移动终端具有随身携带方便、易用等特点，办公人员易将敏感的商业信息和个人资料存入其中，与携带者一起流动，加大了丢失和被窃概率。另外，移动终端的更换周期相对较短，由于缺乏专业的回收或销毁机制，淘汰产品通过二手市场流传，容易造成办公信息泄露。

✓ 移动终端应用程序管理困难

企业应用与个人从第三方市场下载的应用同时并存，未进行有效隔离，一旦某个应用被恶意程序感染，则存在敏感办公信息被泄露的风险。同时后续的升级、新增、下架应用程序等操作，也缺乏统一管理手段。

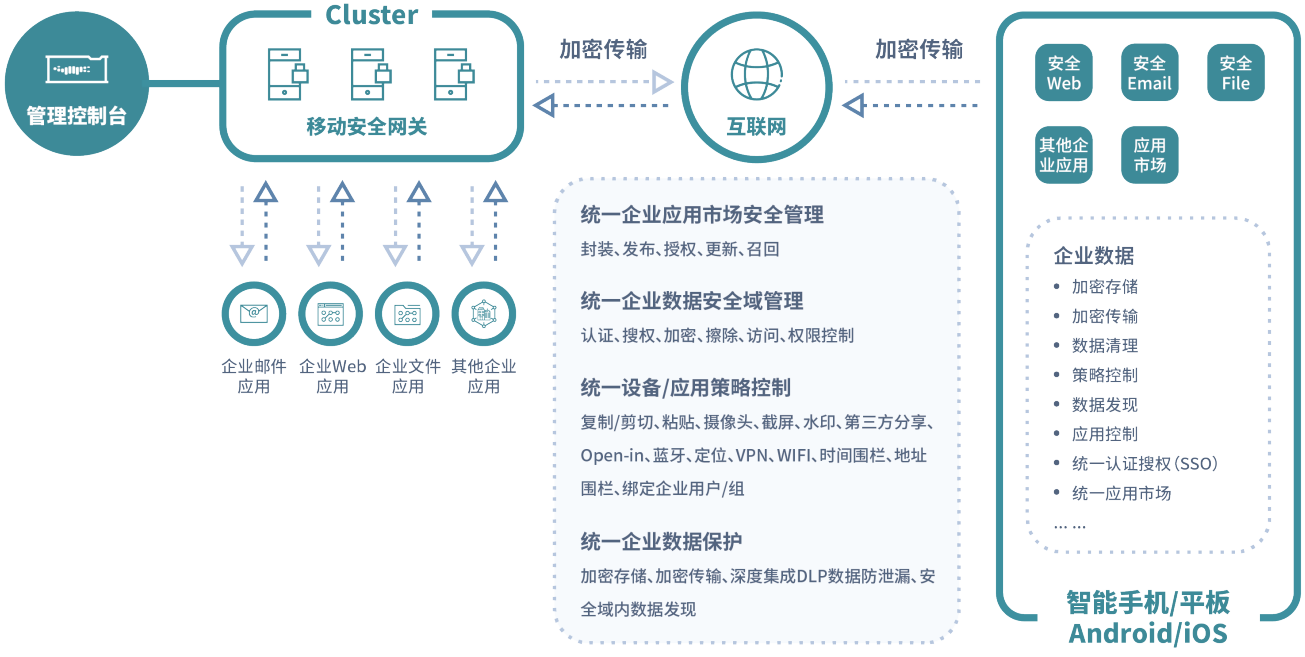
✓ 移动终端上存储未加密的办公数据

若移动终端上的办公数据不进行加密存储，且未与移动终端私人运行环境隔离，一旦被木马、病毒等恶意程序感染，则存在办公系统登录账号、密码被窃取，敏感办公信息被泄露的风险。

✓ 办公数据在移动终端上缺乏隔离措施

普通的移动终端未对办公应用运行环境与个人应用运行环境进行区别对待，但在移动办公场景下，对个人用户而言，容易出现同一终端公私混用的情况。若缺乏有效的隔离措施，容易产生同设备网恶意代码窃取办公敏感数据的风险。

因此，建议组织机构以虚拟安全域技术为核心，从移动设备监控 、企业移动应用管理、移动设备中的企业数据存储安全、移动设备网络传输安全四个角度出发，采用终端安全域作为应用载体，基于统一的数据安全策略平台，实现对移动安全的覆盖。同时，结合其他数据安全产品，为移动应用运行和控制提供统一的数据安全管理环境，实现数据安全有效落地。



b）关键功能

为避免以上安全风险，移动 DLP 需具备以下安全功能：

✓ **身份认证**

移动 DLP 需具有可靠的、唯一的接入身份标识，杜绝身份伪造；确保只有通过身份认证的移动终端能够连接移动办公后台系统，并访问后台数据；确保只有合法用户能够访问移动终端的办公数据。

✓ **链路加密**

通信安全移动终端通过身份认证接入移动办公后台系统后，无线链路需采取可靠的加密措施，以保护办公数据在传输过程中的可用性、完整性、保密性，防止被窃听、被泄露。

✓ **数据加密**

若办公数据在移动终端落地，需采用加密技术进行保护，确保办公数据在移动终端上的完整性、保密性，同时保证办公数据与个人数据隔离存储。

✓ **安全沙箱**

应用运行环境安全。移动办公应用运行时，应采用隔离技术，保证移动终端办公应用与个人应用运行环境有效隔离；终端上多个移动办公应用之间的隔离技术需具有防截屏、防键盘截获等数据防泄露功能。

✓ **内容检测**

移动 DLP 应支持对通过内部应用进行网络外发的数据执行合规性检查，确保用户无法利用网络将敏感数据外发从而导致数据泄露。

✓ **审计安全**

移动 DLP 应支持对设备硬件、网络、系统、应用及用户信息的统计及上报；移动终端应支持安全策略违规事件的记录及上报。

✓ **应用水印**

移动 DLP 应可以要求，对内部应用进行添加水印信息的控制。

✓ **应用管理**

移动 DLP 应可以根据业务需求部署内部应用市场，对单位内部移动应用进行下载、管理、升级等工作。

✓ **数据擦除**

移动 DLP 应支持在管理平台上对丢失的终端进行内部应用的敏感信息的擦除动作，保护敏感信息不被泄露。

✓ **安全控制**

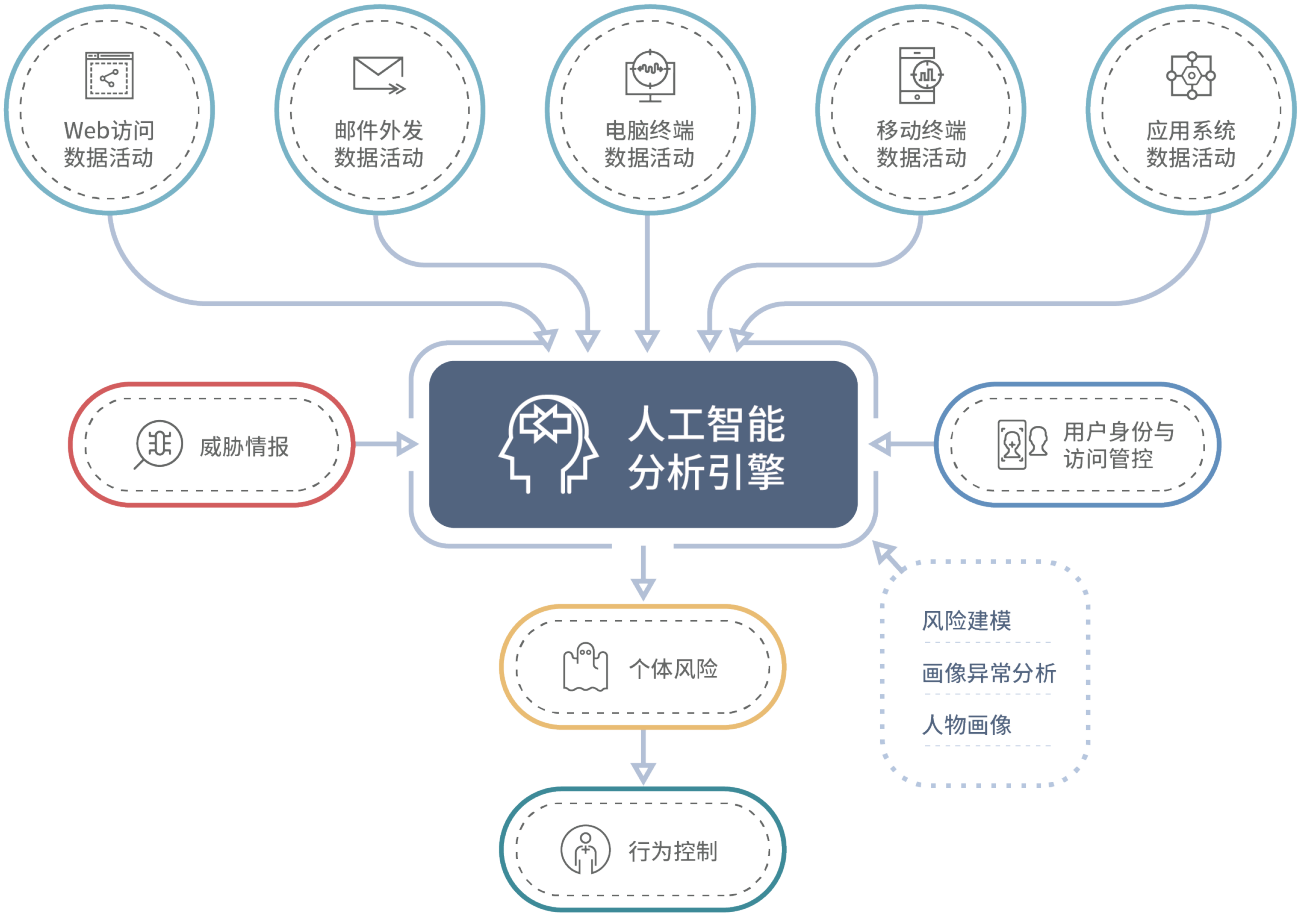
为防止将内部应用中包含的敏感信息通过其他应用向外发送，造成敏感信息的泄露，移动 DLP 应支持当内部应用在移动终端对敏感数据操作行为的阻止和审计，包括复制、粘贴、拷贝、截屏等动作。

六、数据防泄露（DLP）的扩展

1、人工智能与行为分析加入数据防泄露（DLP）体系

关注较多的集中在外部威胁，比如病毒、木马、黑客、DDOS 攻击、APT 攻击等等，近年快速增长且屡屡造成众多企业重大数据安全威胁与泄密的事件更多的开始来自内部威胁。单纯的数据安全防护体系，仅仅关注数据的内容，能够实现数据的可视化、可控等功能，但往往都忽略了数据的另外一个很重要的属性—员工操作行为的风险性，而无法实现对于员工数据操作行为的可视化，也无法从大量的用户和事件中定位高风险用户、高威胁事件，存在安全保护盲区。

因此，组织需要建立一套将行为分析技术与数据保护体系相结合的主动、持续、自适应的防御体系，在实现数据安全保护功能的基础上，能够根据用户的操作行为，实现用户行为的可视化，判断每个用户的风险等级并快速定位风险用户和高危威胁事件，预测可能会发生的威胁风险事件，防患于未然，从而实现智能化、自动化、高效率的主动防御。



通过人工智能的多维度风险建模匹配，实现对内部用户的行为和意图进行监控和预测。在攻击的早期阶段，通过分析用户的 " 非正常 " 数据流行为来锁定内部的潜在威胁，实现了对内部异常用户行为的主动防御，从而有效的实现了内部威胁行为的风险标记。同时，通过系统预置的攻击或异常向量匹配，精确解析出了某些已被木马病毒控制的实体，完成了潜在内部威胁的可视化，并预警和解决了由此而导致的数据泄露风险。

2、数据安全治理的自动化平台

企业利用数字化转型带来技术价值的方式已经从单体服务向微服务、瀑布式开发向敏捷开发、IT 向 DevOps、硬件向基础设施即服务，以及数据中心向公有云发生了一系列的转变 , 这些构成了数字化转型颠覆传统技术革命的基础。

随着 IT 技术的演变 , 企业的安全决策者必须重新规划在 " 新的世界 " 中的数据安全治理方式。原先企业采用的防火墙、IPS 和端点控制这类标准的控制措施已经无法在快速发展的云世界中发挥作用。如今云供应商负责基础架构 , 对于企业而言新的数据控制点已经变成了以身份、数据和业务本身为核心 , 围绕着数据流转使用的每一个环节。围绕着企业在数据安全的现状及痛点 , 越来越多的数据安全自动化相关的处理工具，包括数据建模、数据分类分级、数据识别等自动化辅助工具将协助企业在建立数据安全治理的制度后 , 将制度更有效地实施。系统通过自动化的工作流 , 将不同的数据安全技术工具实践结合在一起 , 为数据安全治理提供了一个完整、可落地的数据安全治理解决方案。

✓ **数据全流程方案**

从企业内的资源发现到数据分类分级 , 再到数据安全策略的配置和执行 , 全方位地覆盖数据安全治理周期的所有环节 ;

✓ AWP 自动化理念 (Automate Where Possible)

最大化的使用自动化 , 将需要人工干预的地方降到最少。在人工介入的必要环节通过工具和界面做到智能辅助 ;

✓ **数据分类分级保护**

通过各种网关的数据防泄露 (DLP) 功能保护由内向外的数据 ; 结合分类分级保护服务、DLP API 服务、脱敏 API 服务等 , 保护企业内部的应用数据。

七、术语与定义

- ✓ **OCR（Optical Character Recognition，光学字符识别：**是指电子设备检查图片上的字符，用字符识别方法将形状翻译成计算机文字的过程。
- ✓ **DLP（Data Loss Prevention，数据防泄露）：**是通过一定的技术手段，防止组织内敏感数据或信息资产，以违反安全策略规定的形式流出组织的一种策略。
- ✓ **SMTP（Simple Mail Transfer Protocol，简单邮件传输的协议）：**主要用于系统之间的邮件信息传递，并提供有关来信的通知。
- ✓ **HTTP/HTTPS（Hyper Text Transfer Protocol，超文本传输协议）：**是一个简单的请求 - 响应协议。HTTPS 在 HTTP 的基础上通过传输加密和身份认证保证了传输过程的安全性。
- ✓ **Hadoop：**利用集群进行高速运算和存储的分布式数据库。
- ✓ **IM（Instant Messenger，即时通讯）：**是指能够即时发送和接收互联网消息等的业务。
- ✓ **NFS/SMB 方式：**当前主流异构平台共享文件系统。
- ✓ **IP（Internet Protocol，网际互连协议）：**是 TCP/IP 体系中的网络层协议，为主机提供数据包传输服务。
- ✓ **HTTP Post：**HTTP 请求方法之一，向指定资源提交数据进行处理请求，数据被包含在请求体中。
- ✓ **HTTP Response：**HTTP 响应方法，在接收 HTTP 请求消息后，服务器会返回一个 HTTP 响应消息。
- ✓ **Active Directory：**是微软 Windows Server 负责网络环境中的集中式目录管理服务。
- ✓ **LDAP（Lightweight Directory Access Protocol，轻型目录访问协议）：**轻量级的目录存储协议，通过 IP 协议提供访问控制和维护分布式信息的目录信息。
- ✓ **OA（Office Automation，办公自动化）：**是将办公和计算机技术结合起来的办公方式。
- ✓ **SIEM (Security Information Event Management，安全信息与事件管理）：**收集网络内的主机系统，安全设备和应用程序生成的日志以及事件数据，并在集中平台上进行整理分析。
- ✓ **SOC（Security Operations Center，安全管理平台）：**收集网络内资产的安全信息，通过对收集到的各种安全事件进行深层的分析、统计和关联，及时反映被管理资产的安全情况。
- ✓ **DHCP（Dynamic Host Configuration Protocol，动态主机配置协议）：**局域网的网络协议。使网络环境中的主机动态的获得 IP 地址、Gateway 地址、DNS 服务器地址等信息。
- ✓ **ICAP（Internet Content Adaptation Protocol，图像采集接口）：**用来从一个传感器捕捉图像数据。
- ✓ **USB（Universal Serial Bus，通用串行总线）：**计算机或其他智能设备与外部设备连接的接口技术。
- ✓ **DVD（Digital Video Disc，数字通用光盘）：**高密度数字视频光盘。
- ✓ **CDROM（Compact Disc Read-Only Memory，紧凑型光盘只读储存器）：**只读光盘。