了解智能数据管理的一般可用性

目录

简介

本文档介绍索引文档匹配(IDM)的一般可用性。

概述

IDM是一种先进的DLP数据分类技术，可显着提高组织有效保护包含敏感数据的文档的能力。

使用IDM，组织可以对保存其敏感数据的文档的内容进行索引和设置指纹。通过创建此数据的指纹存储库，我们的Data Loss Prevention(DLP)产品可以在内容评估期间高效地识别完整或部分匹配的文档。

与使用正则表达式和关键字的传统模式匹配相比，IDM具有显着的优势。IDM允许您匹配实际敏感数据，而不是与任何可能类似于敏感数据的内容匹配。这种有针对性的方法减少了低影响的DLP事件数量，使组织能够将安全运营和资源集中在高价值调查上。

IDM与EDM有何不同？

IDM(Indexed Document Match)和EDM(Exact Document Match)的指纹类型有所不同。

EDM专门关注表格数据的指纹识别，表格数据是以表格格式组织的结构化数据。这意味着EDM旨在处理具有特定结构的数据，如数据库或电子表格。例如，组织可以使用EDM为公司信用卡表设置指纹，从而确保仅监控和保护那些公司信用卡。

另一方面，IDM用于为自由格式文档编制索引和设置其指纹，自由格式文档是不使用特定格式的非结构化数据。IDM能够处理未按表状结构组织的文档，例如文本文件、PDF或Word文档，并对它们设置指纹。

综上所述，IDM用于非结构化数据指纹识别，而EDM用于结构化数据指纹识别。

使用IDM的常见用例有哪些？

一些常见的情况包括指纹识别和保护知识产权，如源代码存储库、专利文件或敏感的公司信息（如HR员工表单、公司文档和法律文档）。

IDM是否根据文件或其文本内容生成指纹？

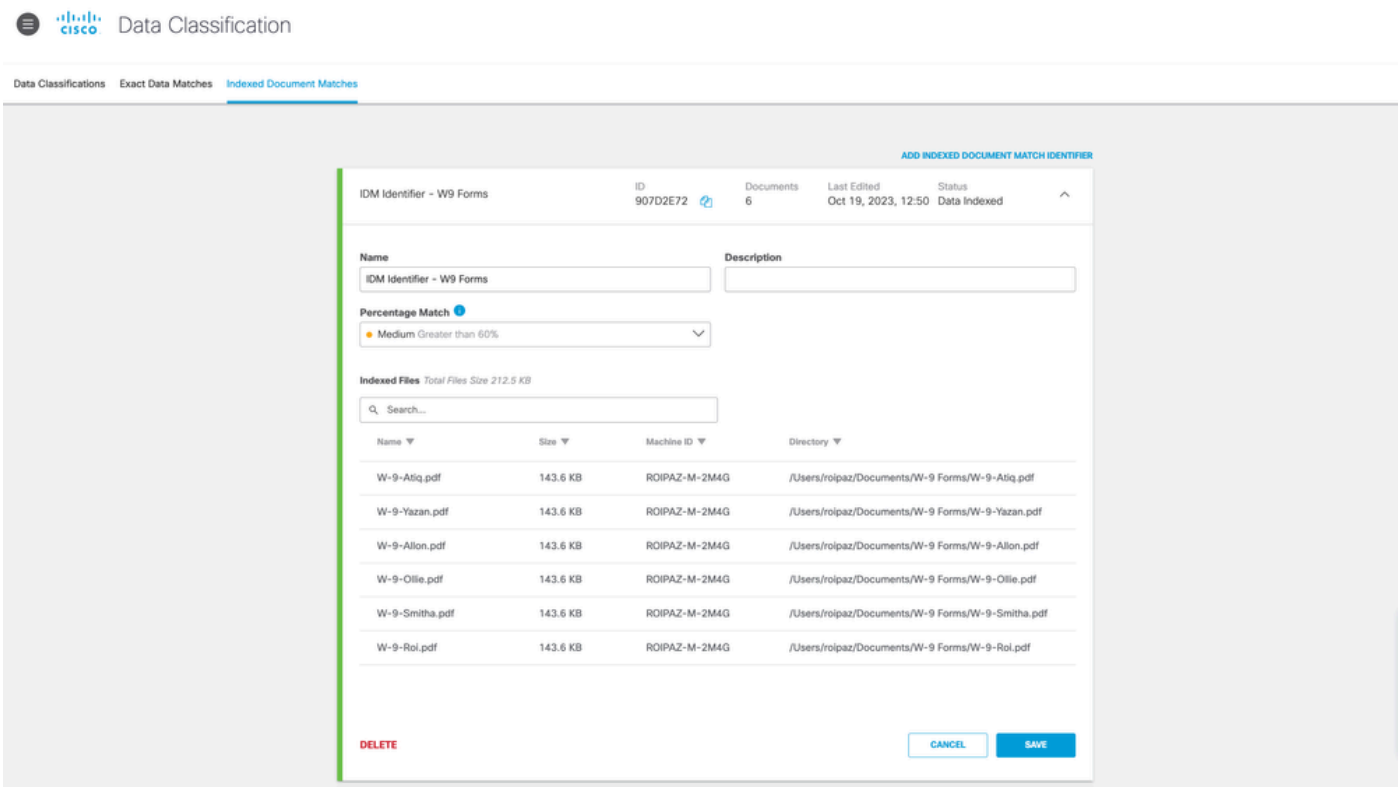
IDM对文档的文本内容而不是文件本身进行索引和指纹。这样，即使某些敏感数据被复制并粘贴到新文件中，IDM也可以与评估的内容进行部分匹配。您可以从预定义选项列表(20%、60%、80%)中进行选择，灵活指定触发违规所需的匹配范围。

如何使用IDM?

Umbrella中的索引文档匹配(IDM)通过生成从敏感文档中提取的文本的散列指纹来运行。然后，多模式DLP的各种扫描会使用这些指纹来完全或部分识别文档的内容。要生成这些指纹，您需要下载并在本地使用思科的DLP索引器工具。

索引器（命令行界面）从文档中提取文本，执行指纹和索引操作，然后散列索引文本。该工具随后将散列指纹上传到Umbrella或Secure Access。

使用索引器工具的输出是在自定义数据分类中使用的新IDM数据标识符类型。这些分类与实时DLP规则和SaaS API DLP规则一起应用，可有效保护静态数据和动态数据。



20327456127636

是否可以安排DLP索引器工具定期指纹识别新数据？

索引器工具可在监控模式下作为后台进程运行。此模式使DLP索引器能够定期自动重新编制索引，确保源数据在Umbrella中定期更新，无需手动操作。

在哪里访问IDM并下载DLP索引器工具？

1. 登录到Umbrella控制面板。
2. 依次导航到策略(Policies)>策略组件(Policy Components)>数据分类(Data Classification)>数据分类(Data Classification)。
3. 点击“索引文档匹配”选项卡。
4. 在本节中，您可以创建IDM标识符并下载DLP索引器。

哪些文件类型与IDM兼容？

IDM支持DLP支持的所有文件类型。您可以在文档中找到支持的文件类型的全面列表。值得一提的是，IDM还支持Unicode字符。

使用IDM时必须考虑哪些限制？

组织中所有IDM数据标识符的索引文本总量不能超过1 GB。当达到分配的配额时，“数据分类”(Data Classification)页面上的“索引文档匹配”(Indexed Document Matches)选项卡会显示警告。

在哪里可以找到更多信息？

Umbrella文档

关于此翻译
思科采用人工翻译与机器翻译相结合的方式将此文档翻译成不同语言，希望全球的用户都能通过各自的语言得到支持性的内容。

请注意：即使是最好的机器翻译，其准确度也不及专业翻译人员的水平。

Cisco Systems, Inc. 对于翻译的准确性不承担任何责任，并建议您总是参考英文原始文档（已提供链接）。