# Comprendre la disponibilité générale d'IDM

## Table des matières


Introduction

Aperçu

- En quoi IDM est-il différent d'EDM ?
- Quels sont les cas d'utilisation courants d'IDM ?
- IDM génère-t-il des empreintes sur la base du fichier ou de son contenu textuel ?
- Comment utiliser IDM ?
- L'outil DLP Indexer Tool peut-il être programmé pour relever régulièrement les empreintes de nouvelles données ?
- Où accéder à IDM et télécharger l'outil DLP Indexer ?
- Quels types de fichiers sont compatibles avec IDM ?
- Quelles limitations doivent être prises en compte lors de l'utilisation d'IDM ?
- Où puis-je trouver plus d'informations ?


## Introduction

Ce document décrit la disponibilité générale de la correspondance de document d'index (IDM).

## Aperçu

L'IDM est une technique avancée de classification des données DLP qui améliore considérablement la capacité de l'entreprise à protéger efficacement les documents contenant des données sensibles.

Grâce à IDM, les entreprises peuvent indexer et relever les empreintes digitales des documents contenant leurs données sensibles. En créant un référentiel d'empreintes digitales de ces données, notre produit Data Loss Prevention (DLP) peut identifier efficacement les documents complets ou partiellement correspondants lors de l'évaluation du contenu.

L'avantage de l'IDM par rapport à la correspondance de modèle traditionnelle utilisant des expressions régulières et des mots clés est significatif. Au lieu d'établir une correspondance avec des données pouvant ressembler à des données sensibles, IDM vous permet d'établir une correspondance avec vos données sensibles réelles. Cette approche ciblée réduit le nombre d'incidents DLP de faible importance et permet aux entreprises de concentrer leurs opérations et ressources de sécurité sur des enquêtes de grande valeur.

### En quoi IDM est-il différent d'EDM ?

IDM (Indexed Document Match) et EDM (Exact Document Match) diffèrent en termes de type de données qu'ils prennent en charge.

L'EDM se concentre spécifiquement sur l'identification des données tabulaires, c'est-à-dire des données structurées organisées sous forme de tableau. Cela signifie que l'EDM est conçu pour gérer des données avec une structure spécifique, telles que des bases de données ou des feuilles de calcul. Par exemple, une entreprise peut utiliser l'EDM pour prendre les empreintes digitales d'une table de cartes de crédit d'entreprise, en s'assurant que seules ces cartes sont surveillées et protégées.

D'autre part, IDM est utilisé pour l'indexation et l'identification des documents de forme libre, qui sont des données non structurées qui n'utilisent pas un format spécifique. IDM est capable de traiter et de prendre les empreintes digitales de documents qui ne sont pas organisés dans une structure de type tableau, tels que des fichiers texte, des PDF ou des documents Word.

En résumé, IDM est utilisé pour l'identification des données non structurées, tandis que EDM est utilisé pour l'identification des données structurées.

## Quels sont les cas d'utilisation courants d'IDM ?

Certains scénarios courants incluent la prise d'empreintes digitales et la protection de la propriété intellectuelle, comme les référentiels de code source, les dépôts de brevets ou les informations d'entreprise sensibles comme les formulaires des employés RH, les documents d'entreprise et les documents juridiques.

## IDM génère-t-il des empreintes sur la base du fichier ou de son contenu textuel ?

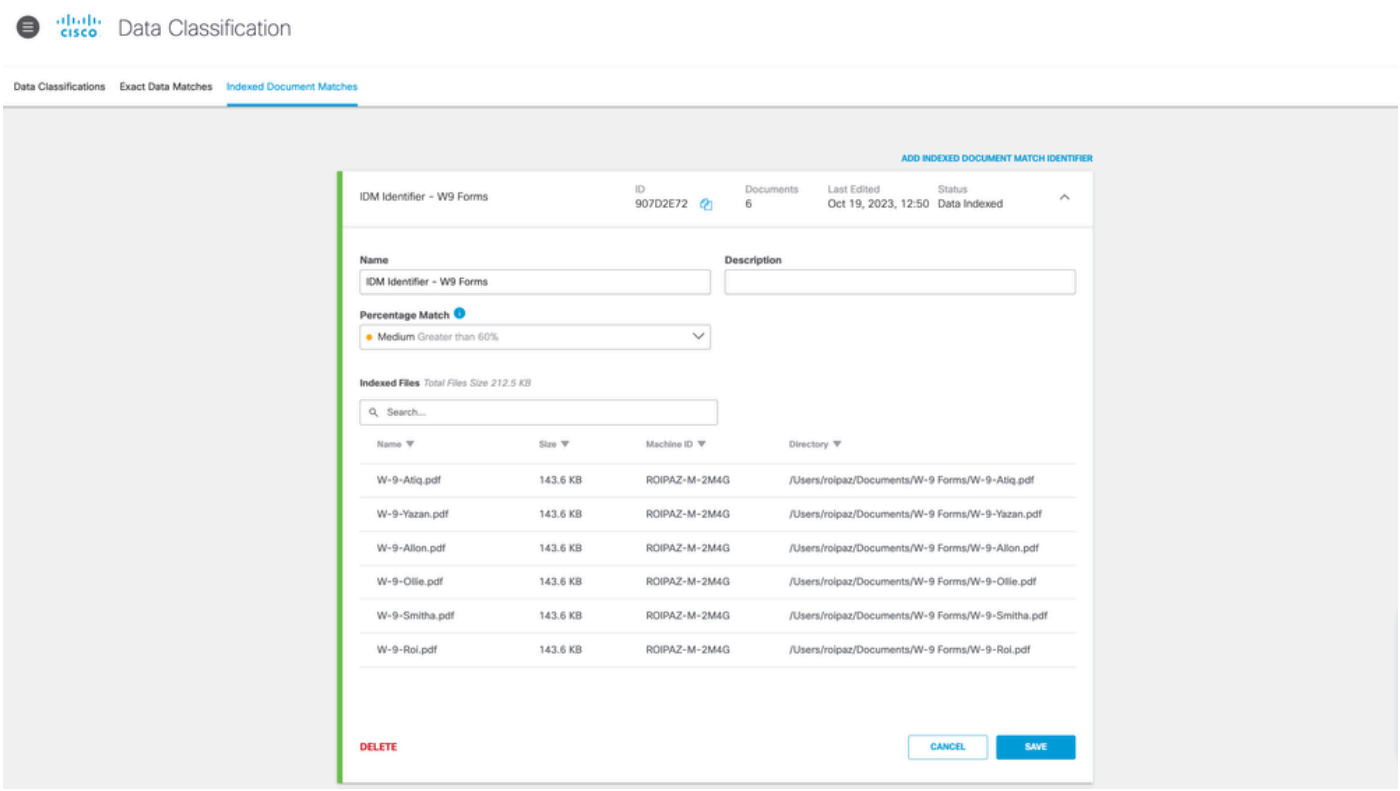
IDM indexe et relève le contenu textuel du document plutôt que le fichier lui-même. Cela permet à IDM d'établir une correspondance partielle avec le contenu évalué, même si certaines données sensibles sont copiées et collées dans un nouveau fichier. Vous avez la possibilité de spécifier l'étendue de la correspondance requise pour déclencher une violation, en sélectionnant une option dans une liste prédéfinie (20 %, 60 %, 80 %).

## Comment utiliser IDM ?

La correspondance de document indexé (IDM) dans Umbrella fonctionne en générant des empreintes de hachage du texte extrait à partir de documents sensibles. Ces empreintes sont ensuite utilisées par les différents balayages de Multi-Mode DLP pour identifier totalement ou partiellement le contenu des documents. Pour générer ces empreintes, vous devez télécharger et utiliser localement l'outil DLP Indexer de Cisco.

L'indexeur, une interface de ligne de commande, extrait le texte des documents, effectue des opérations d'empreinte et d'indexation, puis hache le texte indexé. L'outil télécharge ensuite les empreintes digitales hachées vers Umbrella ou Secure Access.

Le résultat de l'utilisation de l'outil d'indexation est un nouveau type d'identificateur de données IDM à utiliser dans la classification de données personnalisée. Ces classifications sont appliquées avec les règles DLP en temps réel et les règles DLP de l'API SaaS pour protéger efficacement les données au repos et les données en mouvement.




20327456127636

## L'outil DLP Indexer Tool peut-il être programmé pour relever régulièrement les empreintes de nouvelles données ?

L'outil Indexeur peut être exécuté en mode surveillance en tant que processus d'arrière-plan. Ce mode permet à l'indexeur DLP de réindexer automatiquement à intervalles réguliers, garantissant que les données source sont régulièrement mises à jour dans Umbrella sans avoir besoin d'une opération manuelle.

## Où accéder à IDM et télécharger l'outil DLP Indexer ?

1. Connectez-vous au tableau de bord Umbrella.
2. Accédez à Politiques > Composants de la politique > Classification des données > Classification des données.
3. Cliquez sur l'onglet Correspondance du document indexé.
4. Dans cette section, vous pouvez créer des identificateurs IDM et télécharger l'indexeur DLP.

## Quels types de fichiers sont compatibles avec IDM ?

IDM prend en charge tous les types de fichiers pris en charge par DLP. Vous trouverez la liste complète des types de fichiers pris en charge dans la documentation. Il est important de mentionner que IDM prend également en charge les caractères Unicode.

## Quelles limitations doivent être prises en compte lors de l'utilisation d'IDM ?

La quantité totale de texte indexé pour tous les identificateurs de données IDM d'une organisation

ne doit pas dépasser 1 Go. L'onglet Correspondances de documents indexés de la page Classification des données affiche des avertissements lorsque le quota alloué est atteint.

## Où puis-je trouver plus d'informations ?

Documentation générale

À propos de cette traduction
Cisco a traduit ce document en traduction automatisée vérifiée par une personne dans le cadre d'un service mondial permettant à nos utilisateurs d'obtenir le contenu d'assistance dans leur propre langue.

Il convient cependant de noter que même la meilleure traduction automatisée ne sera pas aussi précise que celle fournie par un traducteur professionnel.