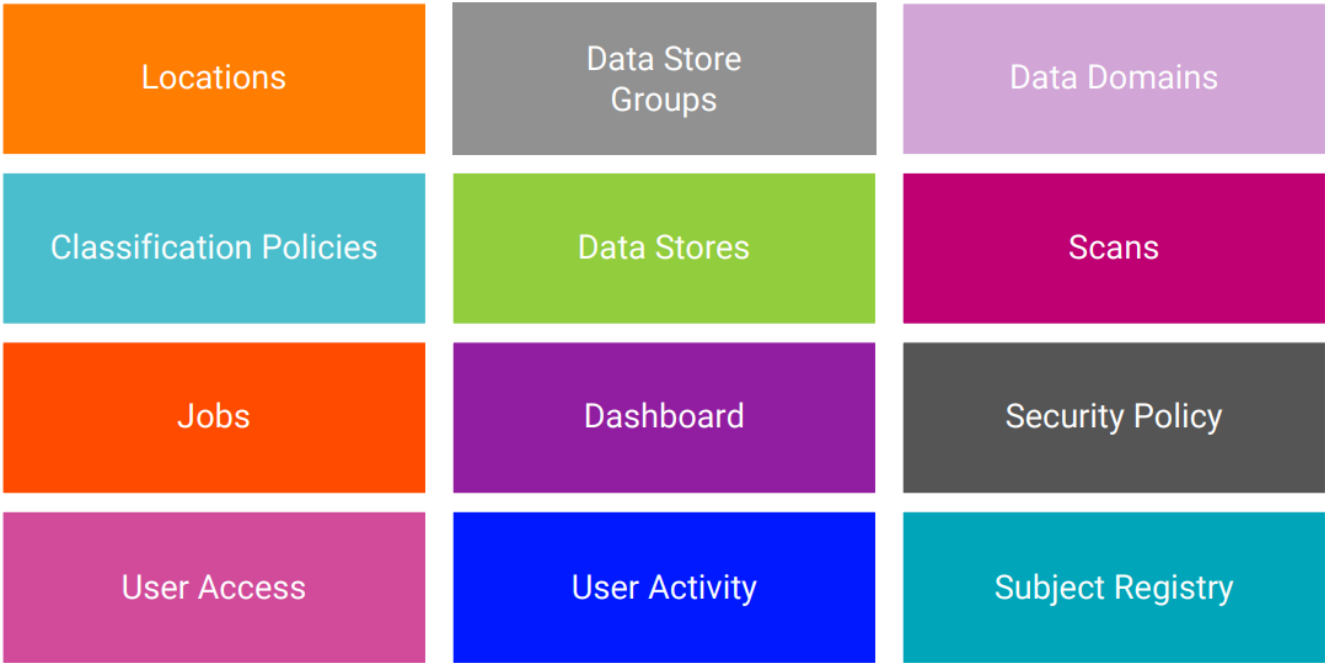


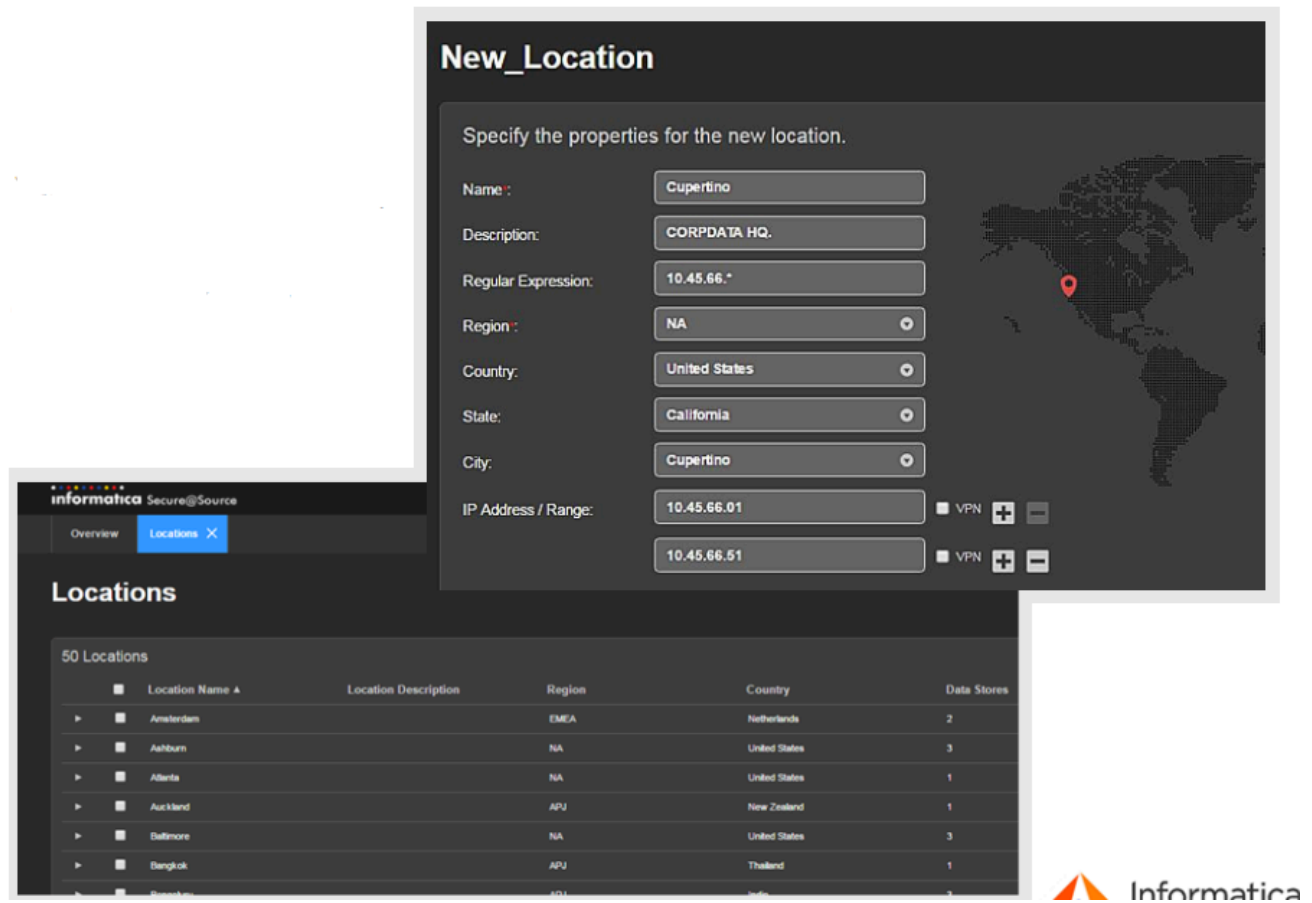
Module 2

Configuration

Core Components



1- Locations



التي تستضيف (Data Centers) هو تمثيل لمنطقة جغرافية تُستخدم لتحديد مكان مراكز البيانات DPM في **Location** (Data Stores). مصادر البيانات

لماذا نستخدم Location؟

- لفهم أين توجد البيانات فعلياً (دولة، مدينة، مركز بيانات).
- لتقسيم مصادر البيانات حسب الموقع.
- لتطبيق سياسات خصوصية مختلفة حسب الموقع الجغرافي.
- لعرض نتائج الفحص (Scan Results) حسب كل Location.

كيف تعمل الـ Location؟

الخطوة	الشرح
1. إنشاء Location	تكتب اسم الدولة/المدينة، وتحدد نمط الـ IP أو Hostname لمصادر البيانات في هذا الموقع.

الخطوة	الشرح
2. ربط Location بمصدر بيانات	الخاص بمصدر Hostname أو IP يحاول يتعرف تلقائيًا على الموقع من خلال DPM البيانات.
3. فحص البيانات	بعد الفحص، تقدر تشوف نتائج الـ Scan لكل Location لوحدها.

مثال عملي:

الحالة:

عندك شركتك فيها 3 مراكز بيانات:

- مصر (Cairo)
- السعودية (Riyadh)
- ألمانيا (Berlin)

خطوات الإنشاء:

1. تنشئ 3 Locations:

- Location: Egypt , Hostname Pattern: *.eg.company.com
- Location: Saudi Arabia , IP Pattern: 192.168.1.*
- Location: Germany , Hostname Pattern: *.de.company.com

2. تضيف Data Store جديد باسم hrdb.riyadh.company.com , Hostname: HR_DB .

3. Saudi Arabia تلقائيًا بـ HR_DB يربط → Location pattern مع hostname يطابق DPM.

4. بعد تشغيل Scan، تقدر تدخل على:

- **Dashboard > Locations**

- وتشوف: "كل نتائج الفحص لمصادر البيانات في Saudi Arabia".

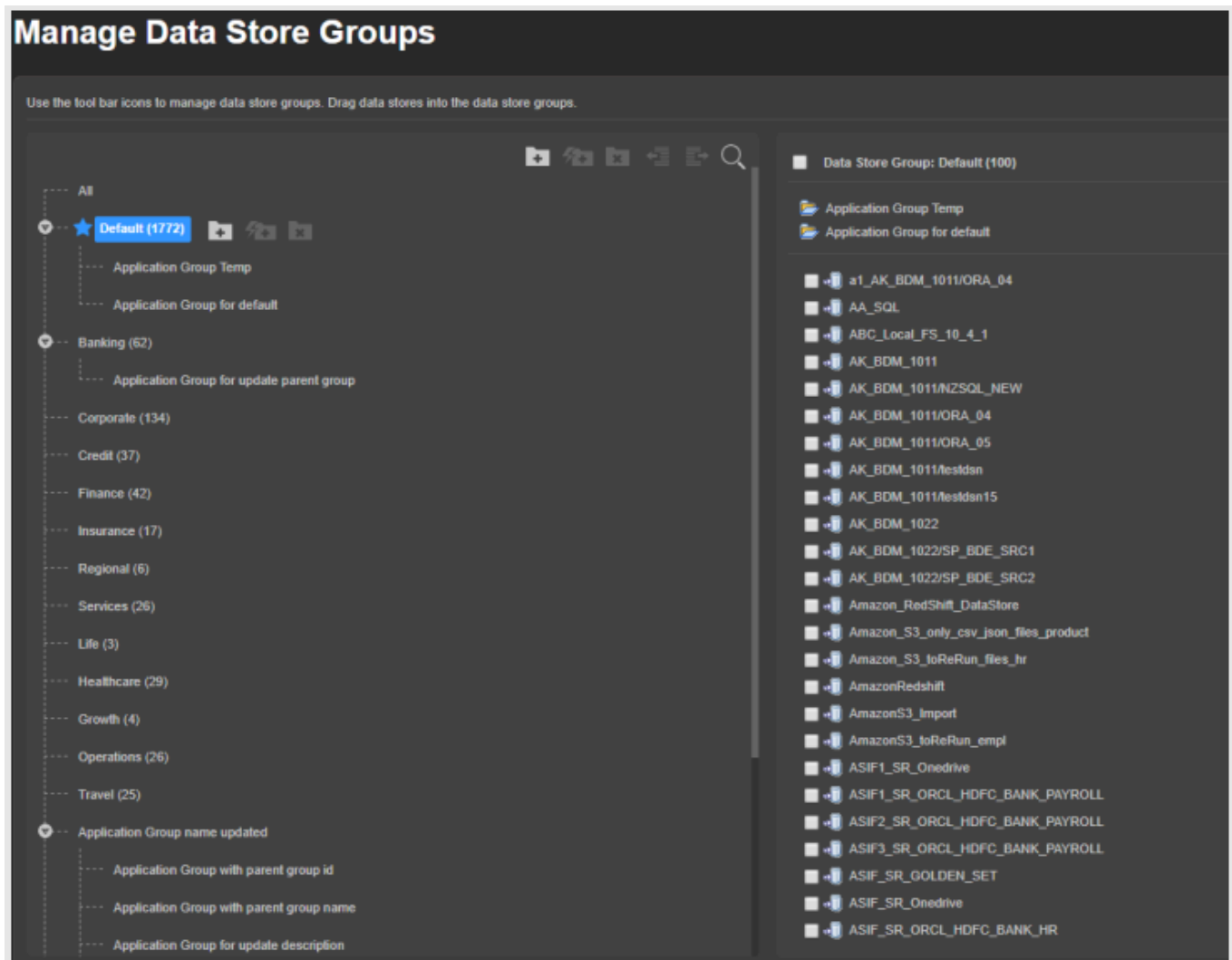
خصائص Location:

الخاصية	الوصف
Name	اسم الموقع الجغرافي (مثلاً: UAE, Germany, ...)
Pattern	نمط IP أو Hostname لتحديد البيانات التي تنتمي للموقع
Assigned Automatically	يربط المصدر بالموقع إذا تحقق الشرط DPM
Manual Assignment	تقدر تعدل يدويًا لو التعرف التلقائي ما اشتغل

فوائد استخدام Locations:

- تنظيم البيانات حسب المناطق.
- تطبيق قوانين خصوصية خاصة بكل دولة.
- تتبع نشاط وفحص البيانات حسب الموقع الجغرافي.
- إعداد تقارير منفصلة لكل دولة أو مركز بيانات.

2- Data Store Groups – Informatica DPM



المرتبطتين ببعض، لتسهيل إدارتهم وتحليلهم سويًا **Data Stores** لعدد من (Group) هو تجميع **Data Store Group**.

لماذا نستخدم **Data Store Groups**؟

- لتجميع قواعد البيانات حسب نوع التطبيق أو القسم.
- لتطبيق سياسات موحدة على مجموعة بيانات مرتبطة.
- لتسهيل تصفية نتائج الفحص (Scan Results) وعرضها حسب كل مجموعة.

أنواع التصنيفات الممكنة:

طريقة التجميع	مثال
حسب نوع التطبيق	Oracle, Salesforce, SAP
حسب خط الأعمال	HR, Finance, Marketing
حسب الإدارة	IT Department, Sales Dept

العلاقة:

Data Store Group → تابع لـ → Data Store

مثال عملي:

تخيل عندك 5 مصادر بيانات:

القسم	النوع	Data Store Name
HR	Oracle	HR_DB
HR	Excel	HR_Reports
Finance	SQL	Finance_DB
Marketing	Salesforce	CRM_SFDC
Marketing	MongoDB	Leads_DB

يمكن تعمل 3 Data Store Groups:

1. HR Group : يحتوي على HR_DB و HR_Reports
2. Finance Group : يحتوي على Finance_DB
3. Marketing Group : يحتوي على CRM_SFDC و Leads_DB

كيف أستخدمهم؟

- لما تشغل Scan، وتدخل تعرض النتائج:
- تقدر تعمل Filter حسب Data Store Group
- تشوف كل النتائج الخاصة بمجموعة معينة بس.

مثال:

Dashboard → Filter by → HR Group

3- Data Domains

Data Domains				
284 Data Domains	284 Data Domains			
1 Not Assigned	<input type="checkbox"/> Data Domain Name ▲	Domain Group	Classification Policies	Data Stores
	<input type="checkbox"/> Account_Status Validates input account information against a user-defined dictionary of v... Last updated on 5/8/2020, 7:01:08 AM by SYSTEM.	SecureAtSourceService	7	2
	<input type="checkbox"/> AccountNumber Validates if the input matches account number. Last updated on 5/8/2020, 7:01:58 AM by SYSTEM.	SecureAtSourceService	8	2
	<input type="checkbox"/> Address Validates if input is an address (supports addresses from US, Canada, ... Last updated on 5/8/2020, 7:01:39 AM by SYSTEM.	SecureAtSourceService	11	26
	<input type="checkbox"/> Admission_date Validates if input is an admission date. Last updated on 5/8/2020, 7:01:47 AM by SYSTEM.	SecureAtSourceService	3	-
	<input type="checkbox"/> Age Validates if the input matches age format. Last updated on 5/8/2020, 7:01:21 AM by SYSTEM.	SecureAtSourceService	14	42
	<input type="checkbox"/> Age_With_Flag Age_With_Flag Last updated on 5/8/2020, 7:03:27 AM by SYSTEM.	SecureAtSourceService	1	6
	<input type="checkbox"/> Age_Without_Flag Age_Without_Flag Last updated on 5/8/2020, 7:02:06 AM by SYSTEM.	SecureAtSourceService	1	6
	<input type="checkbox"/> ALL Last updated on 5/8/2020, 7:01:59 AM by SYSTEM.	SecureAtSourceService	2	100

التي تحتوي على بيانات (Columns) لتحديد الأعمدة DPM هو مجموعة من القواعد التي يستخدمها Data Domain (Data Stores) حساسة داخل مصادر البيانات.

ما الذي يحتويه Data Domain؟

- اسم الدومين (مثال: Email, National ID, SSN)
- قواعد تطابق:
- **Metadata Match:** schema الاسم أو الوصف في الـ.
- **Data Match:** القيم داخل العمود (مثال: رقم يبدأ بـ 01 ويتكوّن من 11 رقم).
- **Proximity Match:** السياق أو القرب من كلمات مفتاحية (مثلاً: عمود بجانبه اسمه "الراتب").

من أين تأتي Data Domains؟

الطريقة	الشرح
Predefined	تأتي مع تثبيت DPM (جاهزة مثل: Email, Phone, CreditCard ...)
Create	تنشئ دومين جديد من واجهة DPM
Import	تستوردهم من ملف CSV

أين تُخزّن؟

- تُخزن داخل **Model Repository Service (MRS)**

من الأقوى؟

:بيستخدم منطق أولوية لتحديد الأعمدة DPM

أفضل تطابق → Metadata + Data Match
 فقط Metadata Match فقط → أقوى من Data Match
 فقط → أقل دقة Metadata Match

القاعدة:

Data overrides Metadata
 Metadata overrides nothing

مثال عملي:

هدفك:

تكتشف كل الأعمدة التي فيها أرقام الضمان الاجتماعي الأمريكي (SSNs).

السيناريو:

• عندك 3 أعمدة في مصدر البيانات:

- SSN
- Social_Security
- Comments مكتوبة ضمن نصوص SSN فيه أرقام ←

إيه اللي بيحصل:

1. Metadata Match:

- يطابق أسماء الأعمدة زي: SSN , Social_Security .

2. Data Match:

- يبحث داخل الأعمدة ويلقي أرقام بالشكل: 6789-45-123 حتى لو كانت في عمود اسمه Comments .

وبالتالي، DPM يحدد الأعمدة الثلاثة إنها تحتوي على بيانات SSN.

ملخص سريع:

خاصية	شرح
الوظيفة	تحديد البيانات الحساسة في الأعمدة
طرق التطابق	Metadata, Data, Proximity
الإنشاء	يدويًا أو استيراد من ملف CSV
التخزين	داخل MRS
الأولوية	Data > Metadata

Data Domains – Column Matching Logic

Update the properties for the data domain.

Name*:

Description:

Column Matching Logic*:

☒ Metadata Match
 ☐ Data Match

☐ In case of a conflict, data overrides metadata condition
 ☒ In case of a conflict, metadata overrides data condition
 ☐ In case of a conflict, data overrides metadata condition
 ☐ Both metadata and data conditions match

Specify a match condition to identify data based on metadata.

☐ Pattern
☐ Reference Table
☐ Rule

ما هو Column Matching Logic؟

هو الإعداد الذي يحدد إزاي DPM يتعامل مع الأعمدة لما يحصل تعارض بين تطابق الاسم (Metadata Match) وتطابق القيمة (Data Match).

الأنواع الثلاثة:

النوع	شرح
Metadata overrides Data	لو العمود اسمه أو وصفه يطابق قاعدة metadata → يتم اختياره حتى لو القيم لا تطابق.
Data overrides Metadata	لو القيم جوا العمود تطابق قاعدة data → يتم اختياره حتى لو الاسم لا يطابق.
Both Metadata & Data	لازم الاسم والقيمة يطابقوا القواعد معًا علشان يتم اختيار العمود.

تفاصيل كل نوع

1. Metadata overrides Data

- التركيز على اسم العمود أو وصفه.
- يستخدم لو أسماء الأعمدة موحدة ومعروفة.
- أسرع في الفحص.
- خطر تجاه false positives لو الأسماء عامة.

مثال:

عمود اسمه SSN ، لكن القيم جواه مش كلها على شكل رقم ضمان اجتماعي → DPM هيختاره.

2. Data overrides Metadata

- يركز على القيم داخل العمود.
- مفيد لو أسماء الأعمدة مش واضحة أو مختلفة.
- أدق لكن أبطأ في التنفيذ.
- يقلل احتمال false positives.

مثال:

عمود اسمه Notes ، لكن فيه قيم شبه أرقام DPM → SSN يختاره.

3. Both Metadata & Data

- الأدق، لكن ممكن يفوت أعمدة صحيحة.
- يستخدم لما تكون عايز تطابق صارم.

مثال:

SSN وقيمه تطابق شكل SSN يختار العمود فقط لو اسمه DPM.

متى تستخدم كل نوع؟

السيناريو	الأفضل
الأعمدة عندك واضحة ومنسقة كويس	Metadata overrides Data
الأعمدة عندك مش ثابتة أو من مصادر خارجية	Data overrides Metadata
عايز دقة شديدة جدًا	Both Metadata & Data

النوع	المعنى
Metadata overrides Data	الاسم يكفي لتحديد العمود
Data overrides Metadata	القيم هي التي تحدد
Both Metadata & Data	لازم الاسم والقيم يطابقوا معاً

Data Domains – Metadata Match Condition

ما هو Metadata Match؟

اللي بيستخدم اسم العمود (أو وصفه) لتحديد إذا (Match Condition) هو نوع من شروط المطابقة **Metadata Match** كان يحتوي على بيانات حساسة.

يتم التحقق من الـ Metadata أثناء خطوة الـ **Profiling** داخل عملية الـ Scan.

طرق مطابقة الـ Metadata

بيسمح بثلاث طرق رئيسية لوصف أسماء الأعمدة DPM:

الطريقة	الشرح
1. Pattern	تستخدم Regular Expressions لوصف نمط اسم العمود.
2. Reference Table	جدول يحتوي على أسماء أعمدة معروفة (رسمية أو بديلة).

الطريقة	الشرح
3. Rule	منطق مركب يحتوي على شروط وتعبيرات لتحديد اسم العمود.

1. Pattern Match

الوصف:

- تعتمد على Regular Expression.
- تستخدم لما تكون أسماء الأعمدة متكررة بنمط معين.

مثال:

- تطابق كل الأعمدة التي اسمها فيه كلمة `(?i).*email` regex CopyEdit `*. email`.

مناسب عندما:

- الأسماء متوقعة أو موحدة.
- عندك naming convention واضح في قواعد البيانات.

2. Reference Table

الوصف:

- جدول يحتوي على قائمة بأسماء أعمدة معروفة لبيانات معينة.
- ممكن يشمل:
- الاسم الرسمي (`EmailAddress`)
- الأسماء البديلة (`email` , `e_mail` , `mail_id`)

مناسب عندما:

- عندك تنوع كبير في كتابة أسماء الأعمدة.
- تتعامل مع أنظمة كثيرة من مصادر مختلفة.

3. Rule

الوصف:

- منطق مركب ممكن يدمج أكثر من شرط (AND / OR).
- أكثر مرونة من الـ Pattern.

مثال:

text

CopyEdit

```
IF column name contains "ssn" OR column name starts with "social" THEN match
```

مناسب عندما:

- تحتاج تطابق أكثر دقة ومعالجة حالات استثنائية.
- عايز تمزج بين الشروط وتتحكم في المنطق.

مقارنة سريعة بين الطرق الثلاث:

الطريقة	المرونة	الدقة	سهولة الإنشاء
Pattern	متوسطة	جيدة	سهولة
Reference Table	عالية	عالية	سهولة نسبيًا
Rule	الأعلى	الأعلى	تتطلب معرفة بالمنطق أو التعبيرات

Data Domains – Data Match Condition

Metadata Match

Data Match

Proximity Match

Specify a match condition to identify data based on the value.

☐ Pattern

☐ Reference Table

☐ Rule

[0-9]{10}

ما هو Data Match؟

اللي بيحدد إن عمود يحتوي على بيانات حساسة بناءً على (Match Condition) هو نوع من شروط المطابقة **Data Match**. القيم داخل العمود (وليس اسمه).

يتم تقييم هذه القيم أثناء خطوة الـ **Profiling** داخل عملية الفحص (Scan).

طرق مطابقة القيم داخل الأعمدة:

الطريقة	الشرح
1. Pattern	تستخدم Regular Expression لتطابق القيم داخل العمود.
2. Reference Table	جدول فيه القيم المقبولة أو المتوقعة لنوع بيانات معين.
3. Rule	منطق مركب يحتوي على شروط لتقييم البيانات داخليًا.

1. Pattern Match

الوصف:

- يستخدم **Regular Expression** لمطابقة القيم داخل الأعمدة.

مثال:

- للتعرف على أرقام SSN: `\d{3}-\d{2}-\d{4}`
- للتعرف على أرقام بطاقات الائتمان: `\d{4}-\d{4}-\d{4}-\d{4}`

مناسب عندما:

- تكون البيانات بتتبع نمط ثابت وواضح.

2. Reference Table

الوصف:

- جدول يحتوي على قائمة بالقيم المتوقعة أو الصيغ الشائعة لنوع معين من البيانات.
- يمكن أن يتضمن:

- كلمات مرور شائعة.
- أسماء دول/مدن.
- أنواع مستندات رسمية.

مناسب عندما:

- بتعامل مع بيانات محدودة أو قائمة ثابتة من القيم المعروفة.

3. Rule (Data Rule)

الوصف:

- منطق شرطي معقد لتقييم القيم داخل العمود.
- يُنشأ في (MRS) Model Repository Service.

مثال:

IF column value is numeric AND length = 16 AND starts with '4' THEN match (Visa card number)

مناسب عندما:

- عايز تتحقق من شروط متعددة معقدة.
- بتعامل مع بيانات فيها أنماط متغيرة أو تحقق إضافي.

مقارنة سريعة بين طرق Data Match:

الطريقة	الدقة	المرونة	سهولة الاستخدام
Pattern	متوسطة-عالية	مرنة	متوسطة
Reference Table	عالية	أقل مرونة	سهلة
Rule	الأعلى	الأعلى	تحتاج معرفة بالمنطق أو إنشاء قواعد

Data Domains – Data Match: Conformance Score

ما هو Conformance Score؟

أثناء خطوة **Data Domain** هو النسبة المئوية للقيم داخل عمود معين التي تطابق قاعدة الـ **Conformance Score** **Profiling** (Database Scan) في فحص قواعد البيانات.

الهدف من Conformance Score

- تحديد إذا كان العمود يحتوي بالفعل على بيانات حساسة.
- بناءً على النسبة المئوية المطابقة، يقرر DPM ما إذا كان العمود:
 - غير حساس.
 - في النطاق الرمادي (Validation Range).
 - حساس فعلاً.

كيفية الحساب

text

CopyEdit

$$\text{Conformance Score} = 100\% \times (\text{عدد القيم المطابقة} / \text{إجمالي عدد القيم})$$

يمكن استثناء القيم **Null** من الحساب لو حبيت.

النتائج المحتملة بناءً على النسبة:

النسبة المتوقعة المطابقة	الإجراء
منخفضة جدًا	يتم رفض العمود كمطابق للدومين
متوسطة (نطاق رمادي)	يتم تسجيله في تقرير التحقق (Validation Report)
عالية جدًا	يتم تأكيد أن العمود يحتوي بيانات حساسة

مثال عملي

عمود: comments

البيانات
123-45-6789
987-65-4321
نص عشوائي
444-55-6666
(null)

إذا كانت قاعدة الـ Domain هي SSN (Social Security Number):

- القيم المطابقة: 3
- القيم الكلية (باستثناء null): 4
- $\text{Conformance Score} = (3/4) \times 100\% = 75\%$

لو إعدادك للدومين محدد إن أقل نسبة مطابقة للحساسية هي 70%، يتم: ☒ اعتبار العمود comments كحساس.

إعدادات إضافية:

- يمكن ضبط النسبة المطلوبة في إعدادات الدومين.
- يمكن تحديد ما إذا كان يتم استثناء القيم الفارغة (null) أو احتسابها.

Classification Policies – Informatica DPM

The image shows two overlapping screenshots from the Informatica Data Protection Manager (DPM) interface.

The top screenshot, titled "New_Classification_Policy", shows a two-step process:

- Specify Properties and Assign Data Domains:** This step includes fields for:
 - Name:** New_Classification_Policy
 - Description:** (empty field)
 - Cost for Each Impression:** 50
 - Sensitivity Level:** (dropdown menu)
- Data Domain Match Conditions:** This step shows a table of data domains assigned to the policy.

The bottom screenshot, titled "Classification Policies", shows a list of 117 policies. A blue box indicates "284 Assign Data Domains". Below this, a table lists several policies:

Classification Policy Name	Data Domains	Data Stores	Scans
ALL			
Policy for quickscan	1	45	30
alphanumeric_underscores_policy			
Policy for quickscan	1	52	52
ASIF_SR_HDFC_CORP_POLICY			
Policy for Subject Registry HDFC Corp	12	5	5
AZURE_POLICY			
Policy for quickscan	1	5	5
AzureDataLake_CP			
Policy for quickscan	1	-	-

ما هي Classification Policy؟

تُستخدم لتحديد وتصنيف البيانات الحساسة بناءً على معايير **Data Domains** هي مجموعة من **Classification Policy**. أمان البيانات الخاصة بالصناعة أو المؤسسة.

مكونات وتصميم السياسة

العنصر	الشرح
Policy	تحتوي على عدة Data Domains

العنصر	الشرح
Data Domains	تحدد أنواع البيانات الحساسة (مثل: SSN، Email)
Risk Cost	تكلفة تقديرية لكل حالة كشف بيانات حساسة (مثلاً: \$100 لكل كشف لرقم بطاقة)

أين تُستخدم؟

- يتم ربط سياسة واحدة على الأقل بأي **Scan Job**.
- عند تشغيل Scan، DPM يستخدم السياسة لتحليل وتصنيف الأعمدة في الـ Data Store.

كيف تُدار السياسات؟

من خلال **Policies Workspace** يمكنك:

- إنشاء سياسة جديدة.
- استيراد أو تصدير سياسة (CSV أو XML).
- تعديل أو نسخ أو حذف السياسات.
- استعراض السياسات الموجودة.

التخزين:

- السياسات تُخزن داخل **DPM Repository**.
- يستدعيها من هناك عند التحرير أو أثناء عمليات الفحص DPM.

أنواع السياسات الجاهزة (Out-of-the-box):

السياسة	الشرح
PII	(مثل الاسم، رقم الهوية) Personal Identifiable Information
PCI	(بطاقات الائتمان) Payment Card Information
PHI	(معلومات طبية وشخصية) Personal Health Information

Risk Cost

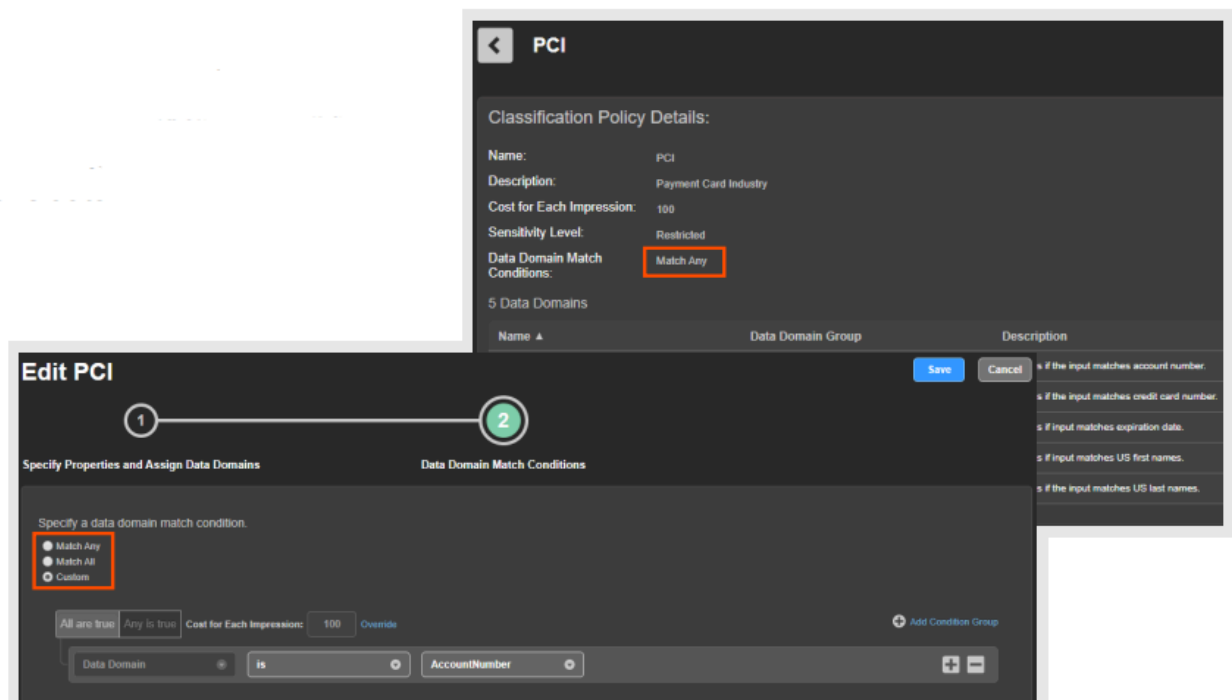
قيمة مالية تقديرية لكل مرة يتم فيها اكتشاف نوع من البيانات الحساسة.

مثال:

- Policy: PCI
- Data Domain: Credit Card Number
- Risk Cost: \$150 لكل ظهور

عند اكتشاف 3 أعمدة تحتوي على أرقام بطاقات → التكلفة التقديرية = \$450

Classification Policies – Data Domain Match Condition



يعني إيه Classification Policy؟

سياسة التصنيف هي مجموعة قواعد بتقول:

"لو لقيت أعمدة في جدول معين تحتوي على بيانات معينة (زي رقم بطاقة، أو رقم قومي)، اعتبر الجدول ده حساس، وحدد درجة حساسيته."

كل سياسة بتحتوي على:

1. Data Domains

وهي أنواع البيانات الحساسة (زي: رقم الهوية، الإيميل، رقم البطاقة)

2. Match Condition

ودي بتقول: هل يكفي وجود دومين واحد؟ ولا لازم كلهم؟ ولا شرط مخصص؟

3. Sensitivity Level

(مثلاً: Confidential). مستوى الحساسية

4. Cost per Record

تكلفة مالية تقديرية لو حصل تسريب لكل قيمة بيانات حساسة

أنواع Match Condition – بالتفصيل

1. Match Any

لو لقيت أي نوع من البيانات الحساسة اللي في السياسة، خلاص اعتبر الجدول حساس.

مثال:

لو عندك سياسة فيها:

- رقم البطاقة

- CVV

وجيت عملت Scan على جدول فيه عمود فيه بس "رقم بطاقة" → خلاص السياسة تعتبر انطبقت.

2.Match All

لازم كل أنواع البيانات اللي في السياسة تكون موجودة في الجدول عشان السياسة تنطبق.

مثال:

لو السياسة فيها:

- رقم البطاقة

- CVV

بس الجدول فيه رقم بطاقة فقط → السياسة ما تنطبقش.

3.Custom

إنت اللي بتحدد الشرط بنفسك.
مثلاً نقول: لازم رقم البطاقة موجود، وواحد من الاثنين (CVV أو Expiration Date).

مستويات الحساسية (Sensitivity Levels)

دي درجات لتحديد مدى خطورة البيانات، وبتستخدمها المؤسسة عشان تعرف تتعامل معاها إزاي.

المستوى	المعنى
Public	بيانات عامة ممكن تتشاف من غير مشكلة (زي اسم شركة).
Internal	بيانات داخلية مش مفروض تطلع بره الشركة، بس مش خطر كبير (زي تفاصيل موظف عادية).
Confidential	بيانات حساسة لازم تفضل سرية (زي راتب، تقييم أداء).
Restricted	بيانات شديدة السرية، تسريبها ممكن يسبب خسائر كبيرة (زي رقم بطاقة، رقم قومي، بيانات صحية).

Cost per Record

يعني لو البيانات دي اتسربت، المؤسسة ممكن تخسر كام لكل سجل؟

مثال:

رقم بطاقة ائتمان = \$100 خسارة لو اتسرب
لو لقيت 10 أرقام بطاقات → الخسارة الكلية = \$1000

ملخص :

العنصر	شرح
Policy	مجموعة قواعد تصنيف بيانات حساسة
Data Domains	أنواع البيانات الحساسة (بطاقة، SSN، إلخ)
Match Condition	يحدد هل أي دوميين كفاية؟ ولا كلهم؟ ولا بشرط خاص؟
Sensitivity Level	درجة خطورة البيانات (Public – Restricted)
Cost per Record	التكاليف لو البيانات اتسربت

مثال تطبيقي:

العنصر	القيمة
Policy Name	PCI Data Policy
Data Domains	Credit Card Number, CVV
Match Condition	Match All
Sensitivity Level	Restricted
Cost per Record	200\$

لو الجدول فيه عمود "Credit Card Number" وعمود "DPM" → "CVV" يصنف الجدول كـ **Restricted** وكل سجل فيه بيتكلف المؤسسة 200\$ لو اتسرب.

Extensions – DPM

Extensions

هي إعدادات جاهزة (plugins) بتحتوي على معلومات اتصال (connection properties)، وتستخدمها عند تنفيذ إجراءات (Actions) في السياسات الأمنية أو المهام (Tasks).

- بتضبط الاتصال مرة واحدة فقط.
- بعد كده، تقدر تعيد استخدامها في أي سياسة أو مهمة.

الهدف من Extensions:

- تسهيل وإعادة استخدام إعدادات الاتصال (بدون تكرار).
- ربط الإجراءات في السياسات الأمنية أو المهام بوسيلة تنفيذ (زي إرسال إيميل، حماية، تسجيل لوج).

أنواع Extensions

النوع	الوظيفة
Custom	تنفيذ مخصص حسب احتياجك (مثل API خاص).

النوع	الوظيفة
Email	إرسال إشعارات بريد إلكتروني.
Protection	تطبيق حماية على البيانات (مثل: Masking).
Service Management	تنفيذ أوامر تخص الأنظمة مثل تشغيل سيرفر أو عملية.
System Log	كتابة سجلات (Logs) في نظام مركزي لمراقبة الأنشطة.

مثال عملي

السيناريو:

أنت عامل سياسة أمان لمراقبة نشاط المستخدم.

لو حصل نشاط غير طبيعي → عايز يبعث إيميل للمدير.

الخطوات:

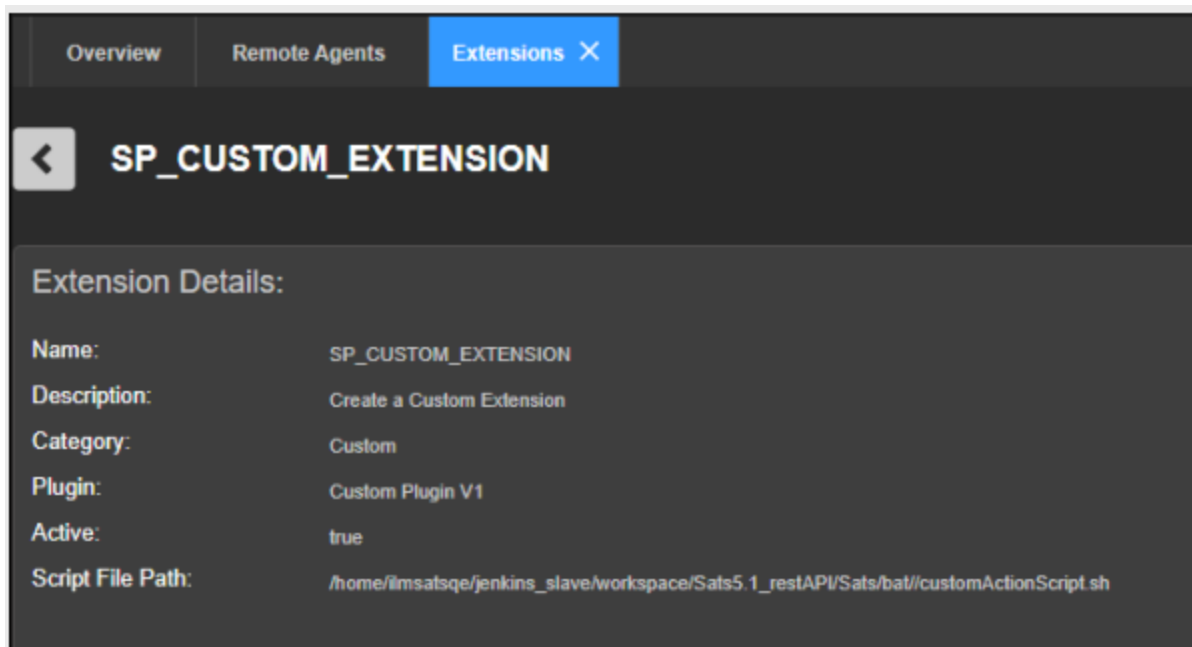
1. تعمل **Email Extension** وتحدد إعدادات SMTP (الإيميل).
2. تحفظ الـ Extension باسم: `NotifyAdminMail`.
3. في السياسة، تضيف Action نوعه "Send Email".
4. تختار الـ `Extension: NotifyAdminMail`.

النتيجة: لما يحصل نشاط مش طبيعي، DPM يبعث إيميل تلقائيًا باستخدام الإعدادات اللي حفظتها.

ملاحظات مهمة

- لازم يكون الـ Extension في حالة **Active** علشان تقدر تستخدمه.
- جاهزة لكل نوع من الأنواع **Plugins** فيه DPM.

Custom Extension – DPM



Custom Extension

بنتيخ لك تنفيذ سكريبت مخصص (مثل .py , .bat , sh على السيرفر أثناء تنفيذ مهمة (Task) أو إجراء (Action) في سياسة أمان.

الهدف منها:

- تنفيذ أوامر خاصة أو إجراءات مخصصة مش موجودة كخيار جاهز في DPM.
- مثال: إرسال إشعار داخلي، تفعيل خدمة، تسجيل في نظام خارجي.

المتطلبات

العنصر	الشرح
Executable File	سكربت تنفيذي (...Shell / Python / Batch)
File Location	لازم يكون على نفس السيرفر اللي عليه DPM
Extension Type	Custom
Plugin	تختار Plugin من النوع: Custom Plugin VI أو DSR Custom Plugin

خطوات الإعداد:

1. اكتب السكريبت اللي عايز تشغله (مثال: notify_team.sh).
 2. ضع السكريبت في مسار على نفس السيرفر اللي فيه DPM.
 3. أنشئ Custom Extension داخل DPM.
 - اختر Plugin: Custom Plugin VI
 - حدد المسار الكامل للسكريبت (مثال: opt/scripts/notify_team.sh/)
 4. استخدم الـ Extension ده في أي Task أو Policy Action.
-

مثال عملي

سكريبت:

bash

```
#!/bin/bash echo "User activity detected" >> /var/log/dpm_alerts.log
```

الخطوات:

- تحفظه في opt/dpm_scripts/log_alert.sh/
 - تنشئ Custom Extension وتحدد هذا المسار
 - تصنيف Action في سياسة أنشطة المستخدم، وتربطه بالـ Extension ده
- النتيجة: لو حصل نشاط معين، DPM يشغل السكريبت → يكتب في اللوج.
-

Email Extension – DPM

EMAIL_EXTENSION_SP

Name:	EMAIL_EXTENSION_SP
Description:	Create a Email Ext with valid username, password
Category:	Email
Plugin:	Email plugin V1
Active:	true
Server Host Name:	mail.informatica.com
Server Port:	25
User Name:	satsemailtest@informatica.onmicrosoft.com
Password:	****
Authentication Enabled:	true
Use Security:	false
Security Protocol:	TLS
Sender Email Address:	satsemailtest@informatica.onmicrosoft.com

هي إعدادات اتصال بريد إلكتروني (SMTP) يستخدمها DPM عشان يقدر بيعت إيميلات تلقائية كجزء من **Tasks** أو **Security Policies**.

- في إجراءات السياسات الأمنية (Security Policies) → مثل إشعار عند نشاط مريب.
- في المهام (Tasks) → مثل إرسال تقرير بعد انتهاء فحص.
- في إجراءات مرتبطة بـ **User Activity Monitoring**.

الإعداد	الشرح
SMTP Server	عنوان خادم الإيميل (مثال: smtp.gmail.com)
Port	رقم البورت (عادة: 465 أو 587)
Sender Email	البريد اللي هُيُرسَل منه

الإعدادات	الشرح
Authentication	اسم المستخدم وكلمة المرور
TLS/SSL	إعدادات الأمان
Recipients (To, CC)	المستلمين

المتاحة Plugins:

- DSR Email Plugin
 - Email Plugin VI
- (تختار واحد منهم لما تنشئ الإكستنشن)

مثال عملي

السيناريو:

فيه سياسة بتراقب المستخدمين، ولما يحصل محاولة وصول غير مصرح → عايزين بيعت إيميل.

الخطوات:

1. تعمل Email Extension:

- SMTP: smtp.office365.com
- Sender: alerts@company.com
- To: security@company.com

2. تسميها مثلاً: SecurityEmailAlert

3. في السياسة الأمنية، تضيف Action نوعه "Send Email"

4. تختار الـ Extension: SecurityEmailAlert

النتيجة: لو حصل خرق → يتم إرسال تنبيه تلقائي بالإيميل.

Protection Extension – DPM

ما هي Protection Extension؟

هي إعدادات بتحدد إزاي **DPM يحمي البيانات الحساسة** عند تنفيذ مهمات الحماية (Protection Tasks).
بتربطها مع **Data Domains** لتطبيق قواعد الحماية تلقائيًا على الأعمدة الحساسة.

الهدف منها:

- تحديد طريقة حماية البيانات الحساسة (مثل الإخفاء، التشفير، التمويه).
- إعادة استخدام نفس إعدادات الحماية في مهام مختلفة.

أين تُستخدم؟

- في **Tasks** المرتبطة بمصدر بيانات (Data Store).
- في **Data Domains**: لتحديد الطريقة الافتراضية لحماية كل نوع بيانات حساس.

أنواع Plugins المتاحة

وظيفة الحماية	Plugin
إخفاء القيم مؤقتًا أثناء الوصول (runtime).	Dynamic Data Masking
تمويه القيم بشكل دائم (تغيير البيانات فعليًا).	Persistent Data Masking
التحكم في صلاحيات الوصول داخل Cloudera.	Cloudera Sentry
التحكم في الصلاحيات داخل Hadoop.	Hortonworks Ranger
تشفير البيانات لحمايتها من التصفح غير المصرح.	Encryption

مثال عملي

السيناريو:

عندك Data Domain اسمه Credit Card Number

وحابب تطبق عليه **Persistent Masking**

الخطوات:

- Plugin: Persistent Data Masking

• تحدد قاعدة الترميز (مثلاً إخفاء أول 12 رقم وترك آخر 4)

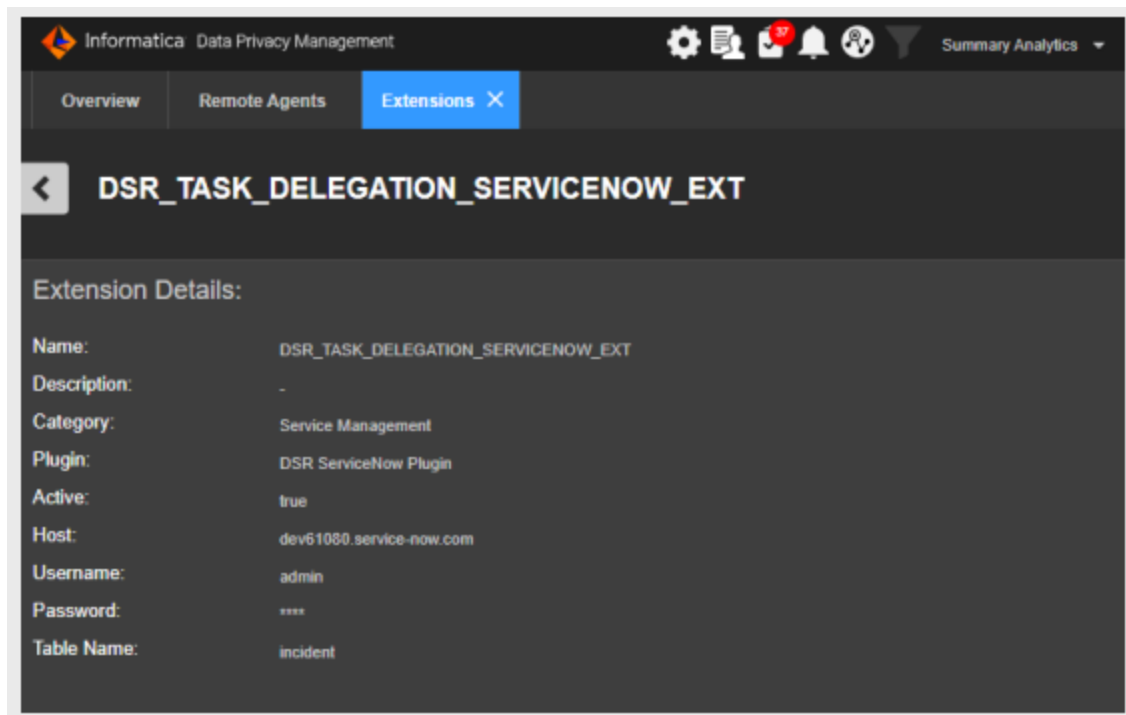
2. تربط الـ Extension بـ domain Credit Card Number

3. تنشئ مهمة حماية (Protection Task) على مصدر بيانات فيه عمود بطاقة

4. تظهر المهمة في Tasks workspace ويتم تنفيذها

إذا العمود موجود في 6 مصادر بيانات، DPM ينشئ 6 مهام تلقائياً – واحدة لكل مصدر.

Service Management Extension – DPM



إنّ عندك أداة DPM بتراقب البيانات الحساسة في شركتك.

لو حصل أي مشكلة (زي حد حاول يدخل على بيانات مش من حقه) أو حد بعث طلب يشوف بياناته (زي GDPR)، إنت محتاج تبلغ فريق تكنولوجيا المعلومات أو الأمن.

بس بدل ما تبلغهم يدويًا، DPM ممكن يعمل ده تلقائي عن طريق إنشاء "تذكرة" (Ticket) في نظام اسمه ServiceNow.

يعني إيه Ticket؟

- زي طلب رسمي بيوصل لفريق معين عشان يبدأ يشتغل عليه.
- زي لما تقول: "في خرق حصل، شوفوه!" → ويوصلهم في البرنامج اللي بيستخدموه.

هنا يلجى دور الـ **Service Management Extension**:

هو اللى بىخلى DPM يقدر بيعت التذكرة تلقائياً إلى ServiceNow.

بيحتوي على إيه؟

- عنوان السيرفر بتاع ServiceNow
- اسم المستخدم والباسورد أو Token
- نوع التذكرة (مثلاً: مشكلة – طلب خصوصية – تنبيه أمني)

مثال:

السيناريو:

1. حصل خرق (واحد دخل على بيانات بطاقة).
2. اكتشف الخرق ده DPM.
3. فيه سياسة (Policy) في DPM فيها إجراء (Action): "اعمل تذكرة".
4. Service Management Extension يروح يشغل DPM.
5. ويعمل تذكرة تلقائية فيها التفاصيل ServiceNow يروح على Extension.

ملخص

العنصر	الشرح
DPM	ببواقب البيانات
Service Management Extension	ببببط DPM مع ServiceNow
النتيجة	أى مشكلة → DPM بيعت تذكرة تلقائياً للفريق المسئول

System Log Extension – DPM



لما أي حاجة بتحصل جوه DPM (زي: مسح بيانات، اكتشاف معلومة حساسة، خرق لسياسة معينة)، بيكون مفيد نسجل الحدث ده في ملف لوج (Log) علشان:

- نرجع له بعدين.
- نحقق في المشكلة لو حصل خرق.
- نتابع النظام من خلال فرق الأمن أو IT.

لكن بدل ما DPM يسجل الحدث ده جواه بس، ممكن بيعته لسيرفر خارجي متخصص بتجميع الأحداث – بنسميه: **System Log Server** أو **SIEM**.

هنا بييجي دور System Log Extension

دي "الإضافة" أو "الإعداد" اللي بتقول لـ DPM:

"لو حصل حدث معين، ابعته كرسالة لوج إلى السيرفر الفلاني."

مكونات System Log Extension

العنصر	الشرح
Plugin	System Log Plugin VI تختار

العنصر	الشرح
Host	عنوان السيرفر اللي هيستقبل اللوجات (مثلاً: 192.168.1.100)
Port	بورت السيرفر (غالبًا 514 لو syslog)
Protocol	نوع الاتصال (TCP أو UDP)
Message Format	محتوى رسالة اللوج اللي هتتبع (ممكن يتضمن: اسم المستخدم – نوع الحدث – الوقت)

مثال

السيناريو:

أنت عامل سياسة في DPM تراقب أي محاولة للوصول لبيانات بطاقات بنكية.

لو حصلت محاولة من شخص غير مسموح له:

- يكتشف ده DPM.
- ينفذ Action داخل السياسة: "Send to System Log"
- أو Splunk أو SIEM زي (تشغل وتبع رسالة لوج إلى سيرفر مركزي خاص بالأمان System Log Extension أو ELK).

الرسالة اللي تتبع مثلاً:

```
[2025-07-09 13:23:10] ALERT: Unauthorized access to CreditCardNumber column by user 'test_user' from IP 10.0.0.55
```

فين ممكن تروح الرسائل دي؟

الرسائل اللي DPM يبعها عن طريق System Log Extension تروح إلى:

- **Syslog Server** (Linux/Unix)
- **SIEM Platforms:** IBM QRadar, Splunk, ArcSight, أو ELK Stack
- **Security Dashboard:** علشان المحللين يقدروا يتصرفوا فوراً

خطوات إنشاء System Log Extension (ببساطة):

1. تدخل على DPM

2. تروح على إعدادات Extensions

3. تعمل **Extension** جديدة

النوع: System Log

- Plugin: System Log Plugin VI

• تدخل عنوان السيرفر (IP أو hostname)

• تختار البورت (مثلاً: 514)

• تضيف إعدادات الرسالة

4. تروح على السياسة، وتضيف **Action**:

• نوعه: "Send to System Log"

• وتربطه بالـ Extension اللي عملته

النتيجة

أي سياسة أو مهمة في DPM بتنفذ الإجراء ده، هتبعث الحدث للسيرفر.

Remote Agent

تخيل إن عندك أداة DPM شغالة على سيرفر مركزي في القاهرة. لكن فيه بيانات موجودة في أماكن تانية:

• مثلاً ملفات على سيرفر في السعودية.

• أو بيانات مشفرة في نظام Hadoop (Cloudera Hive) في دبي.

وهو ببساطة برنامج صغير — **Remote Agent** مش دايماً يقدر يدخل على كل الأماكن دي مباشرة. علشان كده بنستخدم DPM **DPM. بنثبته في المكان البعيد** (زي دبي أو السعودية) علشان يتعامل مع البيانات اللي هناك نيابة عن

الأمان: Two-Way SSL Authentication

الاتصال بين DPM و Remote Agent بيكون مؤمن جداً:

• لازم كل طرف يقدم شهادة (SSL Certificate).

• يعني DPM يثق في الـ Agent، والـ Agent يثق في DPM.

• زي ما بتفتح حساب بنكي وبيطلبوا بطاقتك وبصمتك في نفس الوقت.


1. Protection Remote Agent

بيشغل مع بيانات مشفرة (زي Hive في Hadoop) مهمته يفك تشفير البيانات لما يكون مسموح بناءً على السياسات.

مثال عملي:

- عندك قاعدة بيانات Hive فيها عمود اسمه `credit_card_number` وهو مشفر.
- عامل سياسة في DPM: "بس الناس في قسم الأمن يقدرُوا يشوفوا الرقم الحقيقي"
- لو حد حاول يقرأ العمود:
- الـ Agent يفك التشفير ويرجع القيمة
- لو الشخص مش مسموح له، ما يشوفش القيمة.

- Protection Agent يتواصل مع DPM

 **المهم:** التشفير وفك التشفير مش بيحصل جوه DPM، لكن بيحصل في المكان اللي فيه البيانات علشان الأمان أعلى.

2. Subject Registry Remote Agent

بيشغل مع ملفات غير منظمة (زي Word, PDF, Text...) هدفه يكتشف فيها بيانات الأشخاص (اسم - رقم قومي - بريد إلكتروني...)

مثال عملي:

- عندك ملفات في سيرفر في جده فيها عقود عمل، شكاوى موظفين،... الخ.
- الملفات دي مش قاعدة بيانات، لكن فيها معلومات حساسة عن أشخاص.
- بتنزل Agent هناك، و DPM يقول له:
- "شوفي في الملفات دي أي حاجة تخص شخص اسمه أحمد محمد، أو فيها أرقام قومية"
- الـ Agent يبدأ يفحص الملفات
- يبني فهرس وخريطة هوية توضح البيانات دي فين وبتتكرر فين.

إدارة الـ Agents:

- تقدر تعمل تصدير (Export) لإعدادات الـ Agent إلى ملف CSV.
- تعدّل عليه براحتك (مثلاً تغيّر الـ IP أو البورت).
- وتعمل استيراد (Import) ثاني للملف.

- تقدر كمان تربط كل Agent بمصدر بيانات (Data Store) واحد أو أكثر.

1. Protection Remote Agents
2. Subject Registry Remote Agents

العنصر	Protection Remote Agent	Subject Registry Remote Agent
الوظيفة	فك تشفير البيانات الحساسة	اكتشاف معلومات الأشخاص في ملفات
نوع البيانات	منظمة (Structured – مثل Hive)	غير منظمة (Unstructured – مثل PDF, TXT)
مكان الاستخدام	Hadoop / Cloudera Hive	ملفات في سيرفرات أو مستودعات غير علائقية
الدور الأساسي	تطبيق سياسة الحماية (Protection Policy)	تنفيذ عمليات البحث عن بيانات أفراد (Subject Discovery)
التشفير	يتعامل مع بيانات مشفرة	يتعامل مع بيانات عادية لكن غير منظمة
النتيجة	يحدد ما إذا كان ممكن فك التشفير بناءً على السياسة	يبنى فهرس (Index) وخريطة هوية (Identity Map)
الإخراج	حماية البيانات حسب السياسة	إمكانية توليد تقارير DSAR

Remote Agents Workspace

ده مكان (صفحة جوه DPM) بتقدر منه تتحكم وتدير كل الـ Remote Agents اللي انت عاملهم، سواء كانوا:

- Protection Remote Agents (اللي بيتعاملوا مع البيانات المشفرة)
- Subject Registry Remote Agents (اللي بيكتشفوا معلومات الأشخاص في الملفات)

إيه اللي هتشوفه في الـ Workspace؟

لما تفتح Remote Agents Workspace، هتشوف جدول فيه:

الحقل	معناه
Name	اسم الـ Agent
Host	الجهاز أو السيرفر اللي متسطب عليه

الحقل	معناه
Type	نوع الـ Agent (Subject Registry أو Protection)
User	ممين اللي أنشأ الـ Agent ده

إيه اللي تقدر تعمله من الـ Workspace؟

مممكن تعمل:

الإجراء	معناه
View	تشوف تفاصيل الـ Agent
Edit	تعديل على إعداداته (زي اسم السيرفر – البورت...)
Delete	تمسحه لو مش محتاجه
Copy	تعمل منه نسخة جديدة
Export	تصدر كل التفاصيل لملف CSV (علشان تعدلهم مرة واحدة)
Import	ترفع ملف CSV بعد التعديل
Test Connection	تتأكد إن الاتصال بين DPM والـ Agent شغال
Publish Data Store Info	تبعث بيانات الربط للـ Agent (زي أسماء الـ Data Stores اللي هيتعامل معاها)

دPM في "Action"؟

الـ Action هو تصرف تلقائي أو يدوي بيتم في النظام بناءً على حاجة بتحصل. يعني لما دPM يكتشف مثلاً إن في خرق لسياسة أمنية أو بيانات حساسة انتشرت بشكل مش طبيعي، ممكن:

- بيعث إيميل تنبيه
- يسجل معلومة في Log
- يشغل Script معين
- ينشئ تذكرة في ServiceNow

كل ده اسمه "Action" — وهو خطوة بيقوم بيها النظام تلقائياً أو يدوياً.

إزاي أستخدم Action؟

تقدر تستخدم الـ Actions في حاجتين:

1. في Security Policy

مثال: لو في سياسة بتقول إن ممنوع موظف يشوف بيانات الرواتب، والموضوع حصل، الـ Action هنا ممكن يكون:

- إرسال إيميل للإدارة
- تسجيل خرق (Violation) في التقرير

2. تشغيل يدوي من الواجهة

يعني لو إنت شايف من الواجهة إن في "Anomaly" أو "بيانات حساسة" في مكان معين، تقدر تدوس "Take Action" وتختار:

- تبعت تنبيه
- تشغل Script
- تسجل حاجة في Log

فين تقدر تعمل Action منه؟

من صفحات معينة في DPM زي:

الصفحة	معناها
Anomaly Detection	لو في نشاط مش طبيعي
Proliferation	لما البيانات الحساسة تنتشر في أكثر من مكان
Security Policy Violations	لو حصل خرق لسياسة
Sensitive Fields	لو تم اكتشاف حقول حساس
Subject Details	لو في معلومات عن شخص معين
Top Data Domains	لو في نوع بيانات حساسة منتشر
Top Data Stores	أكثر مصادر بيانات فيها معلومات حساسة

إيه فائدة "Reusable Actions"؟

يعني إنت ممكن تعمل Action مرة واحدة وتستخدمه في أكثر من سياسة أو موقف.
مثال:

- عملت Action اسمه Send Alert to Admin
- استخدمته في 3 سياسات مختلفة
- > بدل ما تعمله 3 مرات، بتستخدمه مرة واحدة

Placeholders – إيه هي؟

لو Action بيععث إيميل أو بيشغل سكربت، ممكن تستخدم متغيرات (placeholders) زي:

Placeholder	بيتحول إلى...
[UserName]	اسم المستخدم اللي عمل الخرق
[DataStoreName]	اسم مصدر البيانات اللي فيه المشكلة
[DomainName]	اسم نوع البيانات الحساسة

يعني لما تشغل الـ DPM، Action بيبدل المتغيرات دي بقيمها الفعلية.

مثال :

السيناريو:

- فيه جدول اسمه employees
- فيه عمود فيه رواتب الناس
- فيه سياسة أمنية بتقول "ممنوع أي شخص خارج HR يشوف العمود ده"

اللي حصل:

- موظف من قسم ثاني فتح العمود
- اكتشف خرق للسياسة DPM

الـ Action:

- فيه Action معمول اسمه Send HR Alert
- بيرسال إيميل إلى مدير الـ HR ويكتب فيه:

"المستخدم [UserName] فتح عمود حساس في [DataStoreName]"

- "Ahmed Salem" بـ [UserName] بيبدل DPM

Action Types

الـ **Action** هو الإجراء الذي DPM يقوم به بعد ما يحصل حدث معين (زي خرق سياسة أو اكتشاف بيانات حساسة).
وفيه 4 أنواع رئيسية من الـ Actions، كل نوع له وظيفة مختلفة:

1. Email Action

المعنى:

يرسل رسالة بريد إلكتروني بشكل تلقائي عند:

- خرق سياسة أمنية
- اكتشاف بيانات حساسة
- تنفيذ طلب من شخص (DSAR)

مثال:

لو Ahmed فتح عمود فيه أرقام فيزا، DPM يبعث إيميل للـ Admin:

"Violation detected by user Ahmed on column 'credit_card_number' in table HR_DB"

2. Service Management Action

المعنى:

ينشئ تذكرة أو طلب (ticket) في نظام إدارة خارجي زي ServiceNow عند:

- خرق سياسة
- طلب بيانات من شخص (DSAR)
- أي مشكلة تحتاج متابعة

مثال:

فيها تفاصيل المشكلة ومين ServiceNow يكتشف أن بيانات موظف ظهرت في مكان غلط → ينشئ تذكرة في DPM
المسؤول عنها.

3. Custom Action

المعنى:

يشغل سكريبت (script) أو ملف تنفيذي (مثل sh. أو bat) لما يحصل الحدث المطلوب.

مثال:

عند خرق سياسة، يشغل سكريبت يحذف ملف من سيرفر، أو يعطل حساب مستخدم من قاعدة البيانات.

ملف السكريبت لازم يكون موجود على نفس السيرفر اللي عليه DPM.

4. System Log Action

المعنى:

يسجل رسالة في سيرفر سجلات خارجي (Syslog Server) علشان ترصد اللي حصل.

مثال:

log يكتب في الـ DPM:

"ALERT: User 'Ahmed' accessed sensitive field 'Salary' in 'Finance_DB'"

متى تستخدم كل نوع؟









النوع المناسب	الحالة
Email	تنبيه شخص مسؤول
Service Management	متابعة عبر نظام تذاكر
Custom	تشغيل أمر مخصص
System Log	تسجيل للأنظمة الخارجية

Risk Score

رقم يمثّل مستوى الخطورة لمصدر بيانات معين (زي قاعدة بيانات أو جدول) = Risk Score
وبيساعدك تعرف بسرعة إيه أكثر الـ Data Stores خطورة في المؤسسة.

إزاي بيتحسب الـ Risk Score؟

افتراضي (weight) بيحسبه تلقائيًا باستخدام 8 عوامل أساسية، وكل عامل ليه وزن DPM

العامل	الوزن الافتراضي
 Sensitivity Level	15%
 Protection Status	15%
 عدد الحقول الحساسة	7%
 عدد السجلات الحساسة	7%
 عدد الـ Targets	15%
 تكلفة البيانات الحساسة	15%
 وصول المستخدمين	15%
 نشاط المستخدمين	7%

مثال :

نفترض إن عندك Data Store اسمه HR_DB ، ودي كانت تفاصيله:

- حساسية البيانات = Restricted
- مش عليه حماية فعالة
- فيه 5 أعمدة حساسة
- فيه 10,000 سجل حساس
- مربوط بـ 3 تطبيقات
- تكلفة البيانات الحساسة = 50,000\$
- 20 مستخدم يقدروا يدخلوا عليه
- حصلت 3 محاولات دخول مريبة

النهائي من 0 إلى 100 بناءً على الأوزان Risk Score ياخذ كل القيم دي، ويحسب رقم الـ DPM.

تقدر تعدّل إيه؟

من صفحة:

ممکن:

- تغيّر وزن كل عامل حسب الأولوية في شركتك.

مثال: لو بتعتبر "نشاط المستخدم" أهم من "عدد السجلات"، تزود وزنه وتقلل الثاني.

- بعد الحفظ، DPM بيشغل **Job** تلقائي يعيد حساب الـ Risk Score لكل الـ Data Stores بناءً على الأوزان الجديدة.

