



Tot plugin PowerBI

Control de versiones	4
Introducción	5
Servicios disponibles en el plugin	5
Modelo de integración	5
Gobierno activo	5
Credenciales requeridas	5
Gobierno activo	5
Configuración necesaria	6
Modelo de base de datos de seguridad	7
Patrón General de Diseño de Tablas	7
Esquema dim	7
Patrón de diseño de las tablas “dim”	8
CommercialPrograms	8
CommercialBusinessUnits	8
Esquema dm_security	8
Tablas comunes a Seguridad Física y Reducción de Datos	9
DSAs	9
Users	9
UsersDSAs	10
ReportsDashboards	10
Tablas de Seguridad Física	10
ReportsPages	10
ReportsPagesDSAs	10
PhysicalSecurity (view)	10
Tablas de Reducción de Datos	11
Dimensions	11
ReportsDimensions	11
RLSCommercialBusinessUnits	11
RLSCommercialPrograms	11
RLSAIldimensions	11
Vistas Reducción de Datos	12
RLSSecurityDSAsUsersWithAllDimensionMembers	12
RLSSecurityCommercialBusinessUnits	12
RLSSecurityCommercialPrograms	12
Actualización de Datos de Dimensiones	13
Actualización de Dimensiones	13
Revisión Inconsistencias Parametrización Reducción de Datos	13
Validación Valores en RLSAIldimensions y resto de tablas RLS	13

Verificación de que existen Valores Parametrizados en Cada Dimensión, Informe y Contrato	14
Parametrización Inicial	14
Tablas de Dimensiones	14

Control de versiones

Versión	Fecha de modificación	Responsable	Aprobador	Resumen de cambios
1.0	28/10/2022	Anjana Producto	Anjana Producto	Creación del documento

Introducción

Servicios disponibles en el plugin

- Crear informes/páginas → Registra informes con sus páginas y/o dimensiones
- Crear/borrar grupos → Registra el DSA con sus usuarios, informes/páginas/dimensiones que aplican en el mismo. El borrado es un borrado lógico de los registros implicados.
- Crear/borrar usuarios → Sirve para añadir o eliminar(borrado lógico) asociaciones de usuario con el DSA
- Borrado de informes → Realiza un borrado lógico (isDeleted = Y) en los registros del informe y de cualquier permiso asociado por algún DSA (ya sea a nivel de página o reducción de datos)

Modelo de integración

Gobierno activo

Dentro del esquema de seguridad de SqlServer se registran los informes, sus páginas, que dimensiones tienen; los dsa, qué usuarios están asociados a los mismos, sobre qué informes/páginas el dsa opera y sobre qué dimensiones.

Con esta información se permite desde PowerBi hacer las consultas que crea oportunas para aplicar los permisos en tiempo real de visualización.

Credenciales requeridas

Gobierno activo

Para el gobierno activo se necesita una credencial de servicio con los siguientes permisos:

- Permiso VIEW DATABASE STATE que se le otorga al usuario sobre la base de datos.
- Permisos SELECT e INSERT que se le otorgan al usuario sobre las tablas del esquema de seguridad.
- Permiso SELECT que se le otorga al usuario sobre las tablas del esquema donde se encuentran los datos maestros de las dimensiones.

Configuración necesaria

```
server:
  port: 15016

totplugin:
  location: http://totpluginpowerbiserver:15016/plugin/powerbi/api/v1
  server:
    url: http://totserver:15000/tot/
  aris:
    - ari: "anja:totplugin:im:/microsoft/powerbi/powerbi/"
    - ari: "anja:totplugin:create:/microsoft/powerbi/powerbi/"
  connection:
    driver: com.microsoft.sqlserver.jdbc.SQLServerDriver
    url: jdbc:sqlserver://rdbservice:1433;database=example
    user: sa
    password: *****
    securitySchema: dm_security
    dimSchema: dim
  metadata:
    reportPagesField: reportPageList
    reportDimensionField: dimensionList
    reportWorkspaceIdField: workspaceId
    reportWorkspaceNameField: workspaceName
    reportPBIField: reportpbi
    dsaPagesField: reportPages
    dimensionPrefix: dim
    separatorAttribute: _-
```

server:

- **port:** El puerto en el que se va a desplegar el plugin.

connection (apartado con la configuración relativa a la base de datos):

- **url:** URL de la localización de la base de datos de SqlServer con el esquema de seguridad y el esquema de datos maestros.
- **user:** El usuario de la conexión con el que se conectará el plugin a la base de datos.
- **password:** La contraseña de la conexión con el que se conectará el plugin a la base de datos.
- **securitySchema:** Es el nombre del esquema dentro de la base de datos definida en url que contiene la tablas del esquema de seguridad (definidas en la parte del modelo de datos de este documento)
- **dimSchema:** Similar a securitySchema, pero el esquema que contiene la tablas de datos maestras (aquellas con los valores de las dimensiones)

metadata (apartado con la configuración relativa a la interpretación correcta del metadata enviado por anjana):

- **reportPagesField:** El nombre del campo del dataset como se ha configurado en Anjana, que contiene la información relativa a que paginas tiene el informe.

- **reportDimensionField**: El nombre del campo del dataset como se ha configurado en Anjana, que contiene la información relativa a qué dimensiones contiene el informe.
- **reportWorkspaceIdField**: El nombre del campo del dataset como se ha configurado en Anjana, que contiene la información relativa al ID del workspace donde está alojado el informe.
- **reportPBIField**: El nombre del campo del dataset como se ha configurado en Anjana, que contiene la información relativa a cuál es el PBI del informe.
- **dsaPagesField**: El nombre del campo del DSA como se ha configurado en Anjana, que contiene la información relativa a que páginas da acceso el DSA.
- **dimensionPrefix**: El prefijo del nombre de los campos del DSA como se ha configurado en anjana, que contienen los valores de dimensiones que da acceso el DSA (Ejemplo de campos con el prefijo dim configurado: dimCountry, dimRegion, etc)
- **separatorAttribute**: El separador utilizado en Anjana para separar los valores de un campo con múltiples valores (como la lista de dimensiones de un informe)

Modelo de base de datos de seguridad

Patrón General de Diseño de Tablas

Todas las tablas citadas en este punto con independencia de su esquema contienen el siguiente patrón de características:

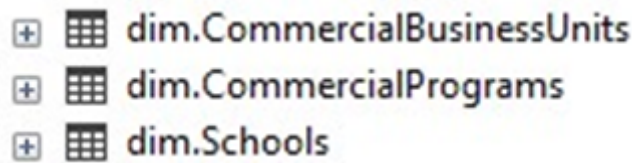
- Tienen primary key que permite identificar unívocamente cada registro
- Las relaciones entre las tablas son valores numéricos para mejorar el rendimiento
- El tipo de borrado a utilizar es lógico (aunque se podrían realizar borrado físico). El borrado lógico se gestiona con un campo llamado IsDeleted que admite los valores Y/N. El motivo para ello es que Anjana también funciona así. Sirve para evitar mostrar datos porque no se consideran vivos pero lo necesitas porque hay información asociada a ella como operaciones históricas, auditoría etc...
- Contiene 4 campos de linaje, dos de creación y dos de última modificación. Uno es el usuario de bbdd que realizó la operación y el otro es la fecha. En el caso de la fecha de modificación junto con la PK permitirá cargas incrementales

```
[IsDeleted]          CHAR (1)          DEFAULT ('N') NULL,  
[CreationRowDate]    DATETIME          DEFAULT (getutcdate()) NOT NULL,  
[CreationRowUser]    NVARCHAR (128)    DEFAULT (user_name()) NULL,  
[LastUpdateRowDate] DATETIME          DEFAULT (getutcdate()) NOT NULL,  
[LastUpdateRowUser] NVARCHAR (128)    DEFAULT (user_name()) NULL,
```

Esquema dim

Este es el esquema referenciado en la configuración de arriba en dimSchema.

Para este caso de uso hemos creado 3 tablas



+ dim.CommercialBusinessUnits
+ dim.CommercialPrograms
+ dim.Schools

Los motivos para la creación de esquema son:

- Contener todas las tablas que se usan para las referencias de Anjana (ej. dim.School). Importante el hecho de que se creen tablas para cargar metadatos de referencia en Anjana no significa que estos datos se carguen/actualicen de manera automática y continua en Anjana. Para eso habría que crear otros plugin a tal efecto
- Contener las tablas con las dimensiones que tienen relación con las tablas de hechos de los datasets usadas en todos los informes de PBI del Cliente. Para el caso de uso que nos ocupa en el informe “Commercial Dashboard V2” se quiere gobernar por parametrización a través de las dimensiones Business Units y Programas, por tanto, para que se den permisos para reducir datos a los data consumers en ese informe necesitaremos tener tablas maestras con los distintos valores posibles de esas dimensiones.

Patrón de diseño de las tablas “dim”

Todas las tablas citadas en este punto además de seguir el patrón general citado anteriormente, deberán seguir las siguientes reglas:

- La primary key se deberá llamar IdDimension
- Deberá contener un campo llamado Name con la descripción del concepto. No podrá ser nulo y su longitud será la que se desee

CommercialPrograms

En esta imagen se ve que la tabla de hechos (Admissions) tiene una relación con la dimensión (bi_keymetricsprogram (programas)).

En la tabla dim.CommercialPrograms se cargarán los datos por los cuales se podrán parametrizar en Anjana contratos que tengan programas. Además, se cargará un registro (-1, Unknown).

CommercialBusinessUnits

Del mismo modo en la tabla dim.CommercialBusinessUnits se cargarán los datos por los cuales se podrán parametrizar en Anjana contratos que tengan business units. En este caso su origen será la tabla bi_macroinformecbd. Además, se cargará un registro (-1, Unknown).

Esquema dm_security

Este es el esquema referenciado en la configuración de arriba en securitySchema.

Tablas comunes a Seguridad Física y Reducción de Datos

En Power BI podemos tener 3 tipos de informes:

- Informes que contienen solo el modelo de datos y los datos (datasets)
- Informes que contienen páginas donde se muestra información, pero no tienen datos (hacen consultas a otros informes de Power BI que contienen datasets)
- Informes que contienen el modelo de datos, datos y páginas (los dos anteriores, esto es poco habitual porque se tiende a modularizar y que cada informe tenga una función)

Seguridad física, en el contexto de los informes, es que si un informe tiene varias pestañas o páginas en un contrato se puede parametrizar un subconjunto de ellas y que las personas que lo firmen solo puedan ver ese subconjunto. La manera de implementar esto es que en la creación de un dataset de tipo informe de Power BI hay que informar las pestañas (páginas) que tiene. En el contrato se mostrará la lista de las pantallas del/los informe/s asociado/s al contrato para que los data steward seleccionen lo que consideren a qué páginas da acceso ese contrato y se guarde en la BBDD. En base a esto como se puede deducir es que la seguridad física aplica a informes que tienen páginas o mixtos (páginas y datos).

La reducción de datos consiste en que las personas pueden ver aquellos datos que tienen parametrizados, por ejemplo, si una persona se le parametriza que puede ver todos los programas de bachelors en las business units de España y Francia solo podrá ver esos datos cuando abra el informe/s donde se aplica esa parametrización. Esa parametrización se guarda en Anjana en contratos y se implementa en Power BI con una funcionalidad llamada "RLS (Row Level Security)". Anteriormente hablábamos de los tipos de informe que nos podemos encontrar en Power BI, pues bien, para hacer contratos con reducción de datos se tienen que hacer para informes que contengan el modelo de datos y los datos o que contengan el modelo de datos, los datos y páginas (mixtos).

Cada informe puede tener su propia parametrización, pero esta debe estar en concordancia con el tipo de informe que es. Si un informe es de tipo dataset puramente, cuando se crea en Anjana no se le añadirán páginas, por lo que no se podrá aplicar seguridad física pero sí reducción de datos. A un informe solo con páginas (sin dataset) se le puede aplicar seguridad física pero no reducción de datos. A los informes mixtos se puede aplicar ambos, seguridad lógica y reducción de datos.

DSAs

Esta tabla contiene información de los DSAs, habrá un registro por cada contrato. Aparte de los campos de linaje y la pk tiene:

- Nombre del contrato
- Id del contrato (generado en Anjana)
- Versión del contrato

Users

Esta tabla contiene información de los usuarios de Anjana. Aparte de los campos de linaje y la pk tiene el campo email que es con el que Power BI gestiona la seguridad a nivel de fila (reducción de datos).

UsersDSAs

Esta tabla contiene un registro por cada usuario asociado a un contrato de Power BI de Anjana.

ReportsDashboards

Esta tabla contiene un registro por cada informe o dashboard de Power BI al que se quiere dar acceso a través de un contrato de Anjana. Esta tabla contendrá un registro por cada report o dashboard del que se comparta un informe. Aparte de los campos de linaje y la pk tiene, del dataset de Anjana le llegará:

- El GUID del report o dashboard
- El nombre del report o dashboard
- El GUID del workspace
- El nombre del workspace que le vendrán del dataset de tipo informe de Power BI
- Versión del dataset

Como se puede ver en la imagen, el GUID del workspace está enmarcado en rojo y en azul el del informe. Para ver esta información basta con abrir el informe en el portal de Power BI u obtenerlo vía API.

Tablas de Seguridad Física

ReportsPages

En Power BI un informe puede tener una o varias páginas, un dashboard solo una. Como su nombre indica esta tabla contendrá las páginas de un informe. Esto es relevante en el sentido de que la seguridad física da acceso a determinadas páginas en función del contrato firmado y eso implica que hay que tener un maestro de páginas por informe. Aparte de los campos de linaje y la pk almacena el Id del report al que está asociado junto con el nombre de la página.

ReportsPagesDSAs

Esta tabla contendrá un registro por cada página de un informe asociado a un contrato de Power BI de Anjana.

PhysicalSecurity (view)

Esta es una vista resumen de todas las tablas relacionadas con la seguridad física. Combina información de los workspaces, informes, páginas de informes, contratos a los que están asociados esas páginas y emails si los contratos han sido firmados por usuarios. Primero se presentan los Ids y luego las descripciones.

Es esta vista la que se expone a Power BI para que se pueda aplicar la seguridad física.

Tablas de Reducción de Datos

Dimensions

Esta tabla contendrá la lista de las dimensiones que contienen los informes. Como se indicó previamente las tablas de dimensiones están en el esquema dim.

Es recomendable crear un proceso que actualice el contenido de esta tabla, si hay nuevas dimensiones las añadirá, si se ha borrado alguna la marcará como borrado (IsDeleted = "Y"). Aparte de los campos de linaje y la pk almacena el nombre de la dimensión.

ReportsDimensions

Esta tabla contendrá la lista de las dimensiones que contiene cada informe. En la creación del dataset se carga esta lista por lo que es en ese momento cuando se carga en esta tabla.

Si en la configuración de anjana el cambio de dimensiones en un dataset no se considera que versiona el dataset es recomendable crear un proceso que actualizará el contenido de esta tabla si se modifica el dataset desde Anjana y lo que se hace es quitar una de sus dimensiones marcará como borrada esa dimensión únicamente (IsDeleted = "Y") en ese informe.

El tener la relación de las dimensiones que aplican a un informe permite para ese informe en la creación de un contrato que se parametrize valores que el data steward o el business translator pueden elegir, ejemplo si a un informe "X" se le asigna la dimensión Business Programs permitirá asociar a un contrato uno o varios o todos los programas.

RLSCommercialBusinessUnits

Esta tabla contendrá la relación de las business units asociadas a un contrato en un informe. Almacenará al menos por cada contrato e informe, una o varias business units.

RLSCommercialPrograms

Esta tabla contendrá la relación de los programas asociados a un contrato en un informe. Almacenará al menos por cada contrato e informe uno o varios programas.

RLSAIIDimensions

Esta tabla contendrá un registro por cada informe en cada contrato que tenga todos los valores de una dimensión asignados. Para parametrizar la reducción de datos de informe en un contrato hay que asignar valores en todas sus dimensiones. Las dimensiones se presentan en listas en la pantalla de Anjana y es el business translator o el data steward quien los selecciona los valores, estos pueden ser pocos, muchos o todos. Si se seleccionan todos una posibilidad era la de coger todos los valores posibles que tiene esa dimensión e insertarlos para ese contrato. Haciendo esto si hay muchos contratos y muchos valores en las dimensiones multiplicaremos el número de registros y el espacio ocupado.

Para evitar esta situación lo que hemos hecho es crear la tabla RLSAIIDimensions donde cada vez que en una dimensión se selecciona en una dimensión de un DSA en lugar de guardarse muchos registros

en la tabla de esa dimensión se guarda un registro en RLSAllDimensions indicando el Id del contrato y el Id de la dimensión.

Posteriormente en Power BI a la hora de realizar la reducción de datos se verifica por cada dimensión si para un usuario:

- Tiene todos los permisos (tiene un registro en la tabla RLSAllDimensions)
- Tiene un subconjunto de valores en la tabla de la dimensión (RLS + nombre dimensión, CommercialProgramas, CommercialBusinessUnits etc)

Un contrato (DSA) tiene que tener para un informe un registro por cada una de sus dimensiones, ya sea en la tabla RLSAllDimensions o en la tabla de la dimensión, en las dos no es posible, no porque no vaya a funcionar sino porque no tiene sentido. Al final el subconjunto de valores de la tabla de la dimensión está incluido en todos los valores de la dimensión. Este ejemplo aclara el comportamiento, para el contrato 3 el data steward en la dimensión CommercialBusinessUnits ha asignado el valor "Spain" y para la dimensión CommercialProgramas Todos. Esto se traduce en:

- 1 registro en la tabla RLSCCommercialBusinessUnits
 - 1 registro en la tabla RLSAllDimensions indicando que tiene todos los programas comerciales
 - 0 registros en la tabla RLSCCommercialProgramas porque tiene asignados todos los programas.
- Es incompatible tener todos los registros de una dimensión y registros en la tabla de dimensión, es una tabla u otra.

Vistas Reducción de Datos

Estas vistas tienen internamente una query que combina todas las tablas relacionadas con la reducción de datos para ofrecer información resumen de lo parametrizado, ofreciendo Ids y descripciones. Estas vistas son las que luego se incorporan al informe de Power BI para hacer reducción de datos por dimensiones.

RLSSecurityDSAsUsersWithAllDimensionMembers

Esta vista indica qué contratos tienen dimensiones donde se han elegido todos los valores. Si el contrato estuviera firmado por varios data consumers, aparecería una línea por cada uno de ellos.

RLSSecurityCommercialBusinessUnits

Esta vista indica qué contratos tienen parametrizado un conjunto de business units. Si el contrato está firmado por varios data consumers aparecerá una línea por cada contrato, business unit y usuario.

RLSSecurityCommercialPrograms

Esta vista indica qué contratos tienen parametrizado un conjunto de programas comerciales. Si el contrato está firmado por varios data consumers aparecerá una línea por cada contrato, programa y usuario.

Actualización de Datos de Dimensiones

En el esquema dim existen tablas de dimensiones que sirven para cargar Metadatos Referenciados en Anjana y una vez ahí para parametrizar contratos.

Estas tablas de dimensiones deberán actualizarse, por cada tabla del esquema dim, contemplando varios casos:

1. Baja de los valores que ya no existen.
 - a. Se conectará al origen de datos y se miembros de la tabla dim que ya no están en el origen de datos, se marcarán como borrados (borrado lógico) haciendo un update en el campo "IsDeleted" = "Y" e informando la fecha de última modificación y usuario de última modificación.
 - b. Importante el registro con valor -1 está excluido de este proceso ya que es un valor incluido a propósito para gestionar que un usuario no vea datos en un informe en el que no tiene firmado un contrato.
 - c. En este caso el proceso deberá informar al equipo de data office de la organización para que informen al business translator/data steward del y expiren los contratos que tengan asociados esos valores borrados
2. Se insertarán aquellos valores que están en el origen de datos pero no existen en la tabla dim

Actualización de Dimensiones

Esta actualización debería realizar las siguientes acciones:

1. Si se han añadido nuevas tablas de dimensión al esquema dim se insertan en la tabla Dimensions un registro por cada tabla. Se creará la tabla RLS{Nombre de la tabla en dim} referenciando a la nueva tabla con los mismo tipos de valores y referencias que el resto de tablas RLS. Generar el nuevo campo en Anjana con los valores de esta nueva dimensión siguiendo las mismas reglas que los otros campos de dimensiones.
2. Se debe actualizar los datos referenciados de Anjana que albergan los valores de las dimensiones existentes se actualizarán desde esta tabla.

```
INSERT INTO dm_security.dimensions ([Object_Id], [DimensionName])
SELECT t.object_id, t.name FROM sys.tables t WHERE t.schema_id = SCHEMA_ID('dim')
AND NOT EXISTS (SELECT * FROM dm_security.Dimensions d WHERE d.object_id = t.object_id)
```

Revisión Inconsistencias Parametrización Reducción de Datos

Validación Valores en RLSAllDimensions y resto de tablas RLS

Se recomienda generar un proceso que revise inconsistencias y valide la integridad de los datos generados en el esquema de seguridad.

Por cada contrato en la tabla RLSAllDimensions y la dimensión asociada, no debe existir en la tabla RLS de la dimensión ningún registro de dicho contrato.

Esto trata de verificar si hay contratos que tienen reducción de datos donde hay parametrizados datos en la tabla RLSAllDimensions y también en la tabla RLS de la dimensión relacionada.

Si hay un registro en la tabla RLSAllDimensions para el contrato X en la dimensión CommercialPrograms no debería haber ningún registro en la tabla RLSCCommercialPrograms para ese contrato X, esto indicaría que existe un bug en el plugin o que alguien ha modificado esos datos en la tabla.

En cualquiera de los dos casos debería ser reportado al equipo encargado del mantenimiento de este sistema..

Verificación de que existen Valores Parametrizados en Cada Dimensión, Informe y Contrato

Por cada contrato se deben llevar a cabo las tareas siguientes

- Recorrer cada dataset de tipo Informe de Power BI
 - Recorrer cada dimensión del informe de Power BI
 - Si hay alguna dimensión que no tiene registros en la tabla RLSAllDimensions y en la tabla RLS de la dimensión correspondiente se debe generar error e informar a alguien ya que todas las dimensiones de un informe en un contrato deben tener valor
 - Tanto si se detecta alguna como si no se pasa a la dimensión, informe o dataset siguiente

Parametrización Inicial

Tablas de Dimensiones

Todas las tablas de dimensiones (esquema dim) es obligatorio que tengan un en su pk (es un campo numérico entero) un registro con valor -1 que representa el valor “Unknown”. Ej. tablas CommercialBusinessUnits y CommercialPrograms.

```
SET IDENTITY_INSERT [dim].[CommercialBusinessUnits] ON
Insert into [dim].[CommercialBusinessUnits] ([IdDimension], [CommercialBusinessUnitId], [Name])
SELECT DISTINCT -1, '00000000-0000-0000-0000-000000000000', 'Unknown'
SET IDENTITY_INSERT [dim].[CommercialBusinessUnits] OFF

SET IDENTITY_INSERT [dim].[CommercialPrograms] ON
Insert into [dim].[CommercialPrograms]([IdDimension], [CommercialProgramId], [Name],
[ProgramShortName])
SELECT DISTINCT -1, '00000000-0000-0000-0000-000000000000', 'Unknown', 'Unknown' SET IDENTITY_INSERT
[dim].[CommercialPrograms] OFF
```

Este valor, aunque no es utilizado por el plugin es necesario para la mecánica de reducción de datos de PowerBI.