

Wer versteckt sich in den Metadaten?

Melanie Hinz
mayato GmbH
Am Borsigturm 9
13507 Berlin
melanie.hinz@mayato.com

Zusammenfassung

Metadaten sind ein wichtiger Bestandteil von SAS. Neben der Management Console oder dem Metadatenbrowser im Window Environment kann man Metadaten unter anderem durch SAS Base Code abfragen oder verändern.

Wir möchten für einen bestimmten Benutzer so viele Information wie möglich anzeigen. Zu welcher Gruppe gehört er? Welche Berechtigungen hat er? Dazu bietet SAS Security Macros an, mit deren Hilfe wir entsprechende Reports erstellen werden.

Schlüsselwörter: Metadaten, Userverwaltung, Report

1 Einleitung

Metadaten oder Metainformationen sind Daten, die Informationen über Merkmale anderer Daten enthalten, aber nicht diese Daten selbst. „Bei den durch Metadaten beschriebenen Daten handelt es sich oft um größere Datensammlungen wie Dokumente, Bücher, Datenbanken oder Dateien. So werden auch Angaben von Eigenschaften eines einzelnen Objektes (beispielsweise Personennamen) als dessen Metadaten bezeichnet.“ [Quelle: Wikipedia]

Metadaten sind das Gerüst von SAS, sie beschreiben Daten und auch die Zusammenhänge verschiedener Metadatenobjekte. Allerdings kann man nur diejenigen Metadaten sehen und auslesen, für die man die Berechtigung hat. Ein vorheriges Backup ist sinnvoll, wenn man die Metadaten nicht nur auslesen, sondern auch modifizieren möchte. Einen guten Einstieg in das Thema bieten [1] und [2]. In [3] erklärt Paul Homes Metadatenreporting mit PROC METADATA.

Metadaten sind strukturierte Daten, die den Ort und die Struktur von Daten beschreiben. Beispiele sind:

- Die Konfiguration verschiedener SAS Server (Workspace Server, Stored Process Server)
- Die Konfiguration verschiedener SAS Produkte (Enterprise Guide, Enterprise Miner)
- User, Gruppen und Rollen
- Zeiger auf von SAS verwendbare Daten (Excel-, CSV-Dateien)
- Spaltenattribute von SAS-Tabellen (Formate, Längen)

2 Wege, Metadaten zu verwalten

Ein erster Weg befindet sich in der SAS Management Console mit dem Plug-In Metadata Manager.

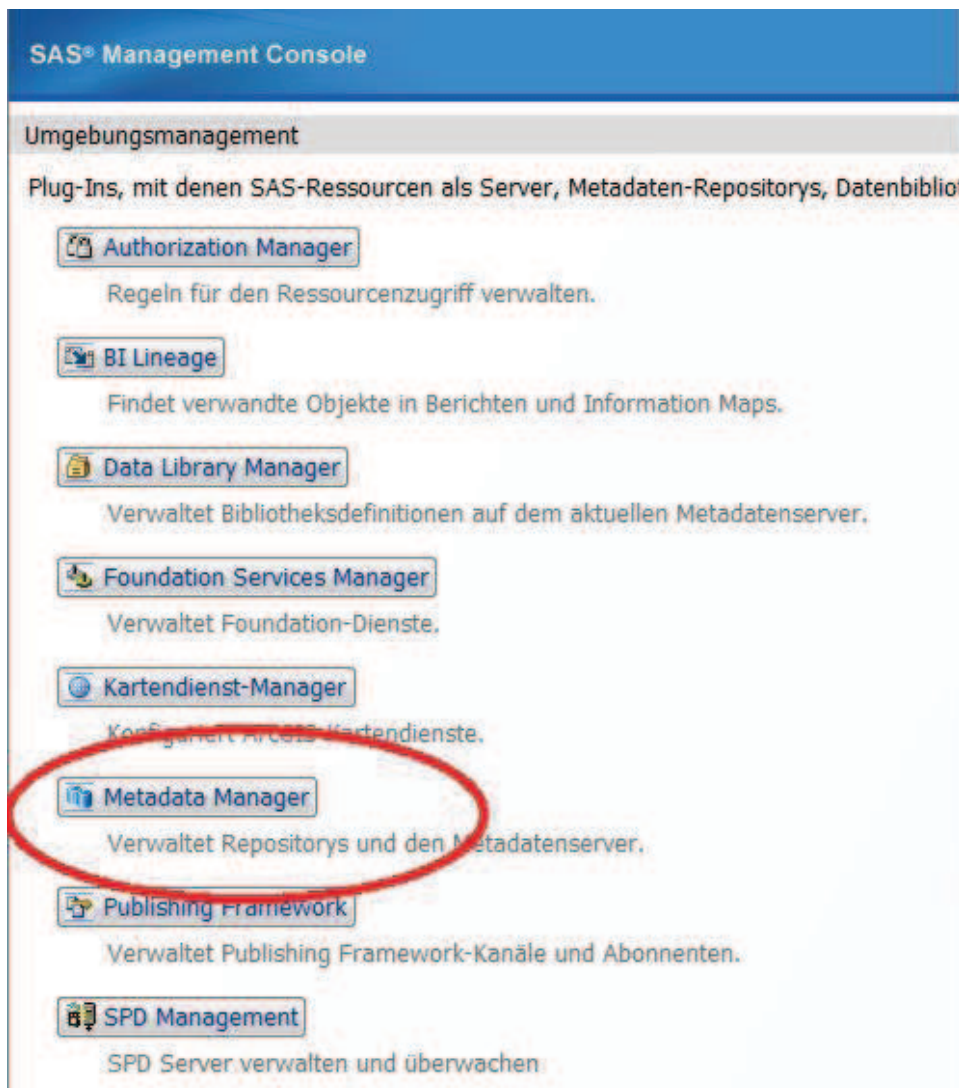


Abbildung 1: Metadata Manager

Um sich mit dem Metadatenserver zu verbinden, benötigt man einige Informationen, wie den Namen des MetadatenServers, die Logindaten, sowie den Namen des Repositorys.

```
options metaserver='metadata-server-hostname'  
        metaport=8561  
        metauser='userid'  
        metapass='password'  
        metarepository='Foundation';
```

Diese Daten gibt man unter anderem beim Metadata Browser im SAS Window Environment ein.

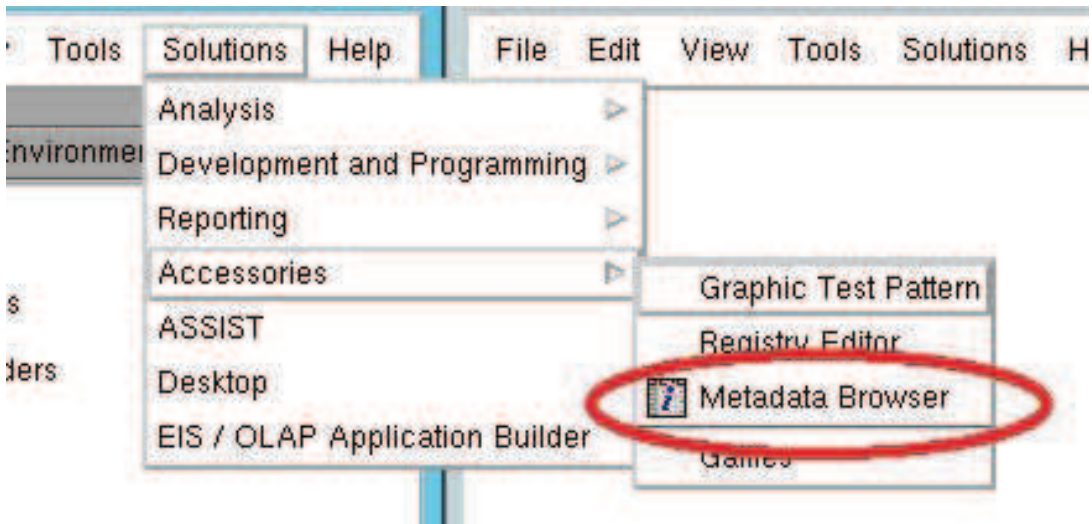


Abbildung 2: Metadata Browser

Dort sieht man auch das zentrale Element der Metadaten, den URI (Unique Resource Identifier). Jedes Metadatenobjekt in SAS ist somit einer eindeutigen Zeichenfolge zugeordnet.

SAS Environment		"Job" Metadatenobjekte			
	Name	Typ	Objekt-ID	Description	Modified(GMT)
mrhelsas94s (Port 8561)	000_PROC_METALIB_Job		A5TUWPB9.C000000S		18. Dezember 2014 21.48 Uhr
BILineage	AB_achttestprogramm (Job		A5TUWPB9.C000001X	Generated by S.	25. Februar 2015 16.49 Uhr
Foundation	AB_neuntesprogramm (Job		A5TUWPB9.C000001Z	Generated by S.	25. Februar 2015 16.49 Uhr
AbstractExtension	Bestuecken_classcopy Job		A5TUWPB9.C000000D		18. Dezember 2014 21.48 Uhr
AbstractJob	DateCheck Job		A5TUWPB9.C000000T		18. Dezember 2014 21.48 Uhr
AbstractPrompt	DESTATIS_Bevoelkeru Job		A5TUWPB9.C000000Z		05. Januar 2015 10.13 Uhr
AbstractProperty	DIFT Organization Dim Job		A5TUWPB9.C000002D		19. März 2015 09.02 Uhr
AbstractTransform	DIFT Populate Order F Job		A5TUWPB9.C000002C		19. März 2015 15.33 Uhr
AccessControl	DM_STUDIE_2015_BA Job		A5TUWPB9.C0000007		18. Dezember 2014 21.48 Uhr
AccessControlEntry	DQ_CARRIAGE_RETU Job		A5TUWPB9.C0000027		10. März 2015 16.58 Uhr
AccessControlTemp	ECC_DATABASE_V2 Job		A5TUWPB9.C000000P		18. Dezember 2014 21.48 Uhr
	ECC_DATABASE_V3 Job		A5TUWPB9.C000000Q		18. Dezember 2014 21.48 Uhr
	ECC_DATABASE Job		A5TUWPB9.C000000O		18. Dezember 2014 21.48 Uhr
	extrakt_testen Job		A5TUWPB9.C0000002		06. Februar 2015 15.19 Uhr

Abbildung 3: Metadata Browser mit URI

Diese URI kann man nun auch mittels PROC METADATA ansprechen. Dazu erzeugt man zuerst einen XML-Request.

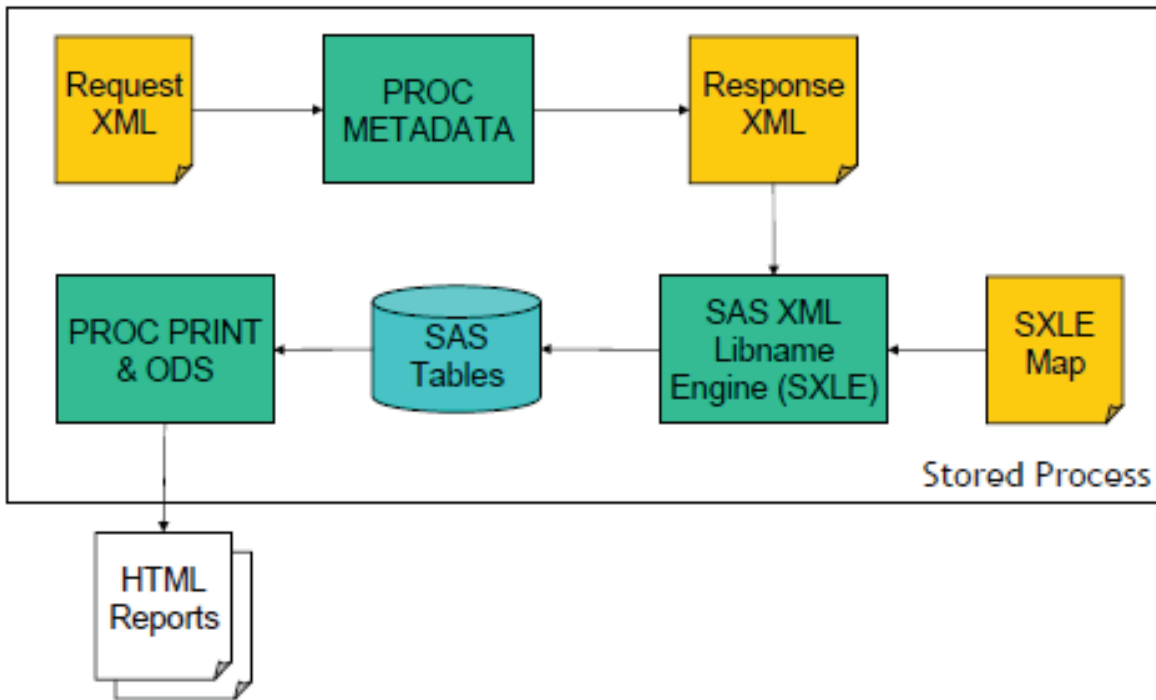


Abbildung 4: XML mit SAS

Im folgenden Code wird eine Anfrage nach der ID und der Emailadresse eines Users gestellt:

```

data _null_;
  file request;
  put '<GetMetadataObjects>';
  put ' <Reposid>$METAREPOSITORY</Reposid>';
  put ' <Type>Person</Type>';
  put ' <Objects/>';
  put ' <NS>SAS</NS>';
  *Flags: + OMI_GET_METADATA(256) + OMI_XMLSELECT(128) +
  OMI_TEMPLATE(4) = 388;
  put ' <Flags>388</Flags>';
  put ' <Options>';
  put " <XMLSelect search=""Person[@Name =: 'Hinz']""/>";
  put ' <Templates>';
  put ' <Person Id="" Name="">';
  put ' <EmailAddresses />';
  put ' </Person>';
  put ' <Email Address="" />';
  put ' </Templates>';
  put ' </Options>';
  put '</GetMetadataObjects>';
run;
  
```

Nun wird PROC METADATA ausgeführt:

```

proc metadata in=request out=response;
run;
  
```

Auf die Anfrage nach Namen und Emailadresse bekommt man die Antwort:

```
<GetMetadataObjects>
  <Reposid>A0000001.A5Y8I8TA</Reposid>
  <Type>Person</Type>
  <Objects>
    <Person Id="A5Y8I8TA.AR000006" Name="Melanie Hinz">
      <EmailAddresses>
        <Email Id="A5Y8I8TA.BT000001" Address="melanie.hinz@mayato.com"/>
      </EmailAddresses>
    </Person>
  </Objects>
<!-- original request deleted -->
</GetMetadataObjects>
```

Mit Hilfe einer XML-Map kann man diese nun nach SAS einlesen und in Tabellen und Berichten weiterverarbeiten.

3 SAS Data Steps, die Grundlagen

Um als Einstieg sämtliche Objektarten zu erhalten, kann man folgenden Code ausführen. Mit `metadata_getntyp` bekommt man dort die n-te Objektart ausgegeben.

```
data _null_;
  length type $64;
  do until (rc<0);
    i+1;
    rc=metadata_getntyp(i,type);
    if rc>0 then put type;
  end;
run;
```

Das Ergebnis ist umfangreich, hier nur ein paar Beispiele:

- Directory: Physischer Ordner
- ExternalTable: Externe Tabelle, z.B. Excel oder CSV
- Job: DI-Studio-Job
- PhysicalTable: SAS-Tabelle
- Transformation: Transformation in einem Job, z.B. Extract-Knoten
- Tree: Ordner im DI-Studio
- Person

Uns interessiert jetzt der User, also die Objektart Person. Mit dem folgenden Code lesen wir die Attribute, sowie deren Werte aus.

```
data ausgabe;
  length attr $64 value $64;
  do until (rc<0);
    n+1;
```


M. Hinz

```
rc=metadata_getnattr("omsobj:Person?@Name='hinzmela'",
                    n,attr,value);
if rc>0 then output;
end;
run;
```

Das Ergebnis ist:

	n	attr	value	rc
1	1	UsageVersion	1000000	12
2	2	Title		12
3	3	PublicType	User	12
4	4	Name	hinzmela	12
5	5	MetadataUpdated	10Mar2015:10:03:39	12
6	6	MetadataCreated	10Mar2015:10:03:39	12
7	7	LockedBy		12
8	8	IsHidden	0	12
9	9	DisplayName	Melanie Hinz	12
10	10	Desc	SAS Entwickler	12
11	11	ChangeState		12
12	12	Id	A5MJGQ2K.AP0004A5	12

Abbildung 5: Attribute eines Users

Diese Informationen reichen uns aber noch nicht. Wir wollen auch wissen, welche Berechtigungen dieser User hat und in welchen Gruppen er ist.

4 Die Security-Makros

Von SAS gibt es vorgefertigte Makros, mit denen wir Informationen aus den Metadaten ziehen können:

%MDUEXTR benutzt PROC METADATA, um Identitätsinformationen zu extrahieren. Wir verbinden uns mit dem Metadatenserver, siehe Kapitel 2. Anschließend wird der folgende Code ausgeführt:

```
%mduextr(libref=work)
```

Die Bibliothek kann dabei eine beliebige SAS-Bibliothek sein. Durch diesen Code werden die Kanonischen Tabellen (canonical tables) erzeugt, die in der angegebenen Bibliothek abgelegt werden.

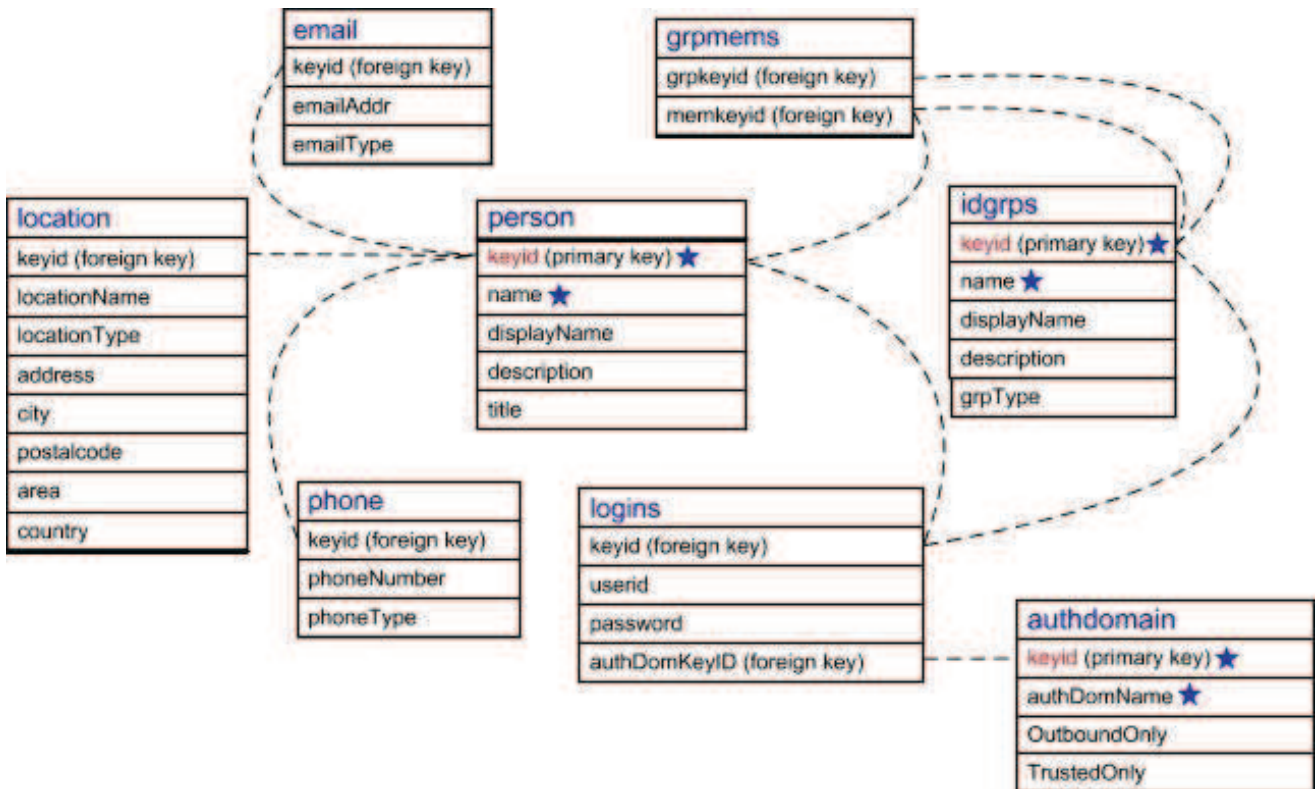


Abbildung 6: Kanonische Tabellen

Diese Tabellen müssen nicht alle befüllt sein. Man legt zum Beispiel nicht immer die E-Mail oder die Telefonnummer eines Users in den Metadaten ab. Pflichtfelder sind mit einem Stern gekennzeichnet. Die Informationen unseres Beispielusers sehen hier so aus:

	keyid	name	description	title	displayname	objid
1	A5MJGQ2K.AP0004A5	hinzmela	SAS Entwickler		Melanie Hinz	A5MJGQ2K.AP0004A5

Abbildung 7: Tabelle PERSON

Wir wollen auch wissen, in welchen Gruppen dieser User ist:

	grpkeyid	memkeyid
1	A5MJGQ2K.A50000XM	A5MJGQ2K.AP0004A5
2	A5MJGQ2K.A50000XR	A5MJGQ2K.AP0004A5
3	A5MJGQ2K.A50000XS	A5MJGQ2K.AP0004A5
4	A5MJGQ2K.A50000XU	A5MJGQ2K.AP0004A5
5	A5MJGQ2K.A50000XW	A5MJGQ2K.AP0004A5
6	A5MJGQ2K.A50000XX	A5MJGQ2K.AP0004A5
7	A5MJGQ2K.A50000XZ	A5MJGQ2K.AP0004A5

Abbildung 8: Tabelle GRPMEMS

Uns interessiert auch, wie die Gruppen überhaupt heißen:

	keyid	name	description	grpType	displayname	objid
1	A5MJGQ2K.A50000XM	CO_Developers			CO: Developers	A5MJGQ2K.A50000XM
2	A5MJGQ2K.A50000XR	DDS_Developers	Entwickler DDS		DDS Developers	A5MJGQ2K.A50000XR
3	A5MJGQ2K.A50000XS	DQDM_Developers			DQDM: Developers	A5MJGQ2K.A50000XS
4	A5MJGQ2K.A50000XU	FIN_Developers			FIN Developers	A5MJGQ2K.A50000XU
5	A5MJGQ2K.A50000XW	ODS_Developers	Entwickler Projekt ODS		ODS Developers	A5MJGQ2K.A50000XW
6	A5MJGQ2K.A50000XX	RAS_Developers			RAS: Developers	A5MJGQ2K.A50000XX
7	A5MJGQ2K.A50000XZ	RWA_Developers			RWA: Developers	A5MJGQ2K.A50000XZ

Abbildung 9: Tabelle IDGRPS

Mit ein wenig Rekursion bekommen wir dann auch die vererbten Gruppen, da eine Gruppe auch wieder Mitglied einer Gruppe sein kann. Dies kann man natürlich nicht nur für einen User machen, sondern sich insgesamt eine Abfrage bauen, die angibt, welche User in welchen Gruppen sind.

Außer den Gruppen interessieren uns noch die Berechtigungen, die unser Beispieluser hat.

```
%mdsecds (identitynames="hinzmela",
          identitytypes="Person",
          folder="\ETL\CP",
          membertypes="Folder",
          includesubfolders=YES);
```

Hier fragen wir nach dem User hinzmela und schränken uns auf einen Ordner ein. Wir betrachten nur die Ordner, inklusive Unterordner. Diese werden aufgelistet. Es wird eine Liste der Ordner ausgegeben, mit Berechtigungen, wie Lesen, Schreiben, Löschen, physisch und auf den Metadaten.

Weiterführende Informationen finden sich in [4].

Literatur

- [1] E. Muriel, Exploring the Metadata Family Tree, SAS Global Forum 2009, Paper 097--2009.
- [2] E. Muriel, P. Simkin, Metadata for SAS 9 Programmers, SAS Global Forum 2008, Paper 134--2008.
- [3] Paul Homes, Metadata reporting with SAS software, TekEds.com
- [4] SAS 9.4 Intelligence Platform: Security Administration Guide