



Was wir vorhersagen, soll auch eintreffen!



SAS Metadatenmanagement Reporting und Analyse //

Ihr Ansprechpartner:

Melanie Hinz
melanie.hinz@mayato.com

mayato GmbH
Am Borsigturm 9
13507 Berlin

www.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. Einige Wege der Zugriffe auf Metadaten werden aufgezeigt. Es werden Beispiele vorgestellt, in denen Metadaten nur mittels SAS Base Code bearbeitet werden.

Schlüsselwörter: Metadaten, Reporting, XML, PROC METADATA

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 in 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)

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

```
options metaserver='metadata-server-  
hostname'
```

Ihr Ansprechpartner:

Melanie Hinz
melanie.hinz@mayato.com

- Spaltenattribute von SAS®-Tabellen (Formate, Längen)

2 Wege, Metadaten zu verwalten

Ein erster Weg findet sich in der SAS Management Console, das Plugin Metadata Manager

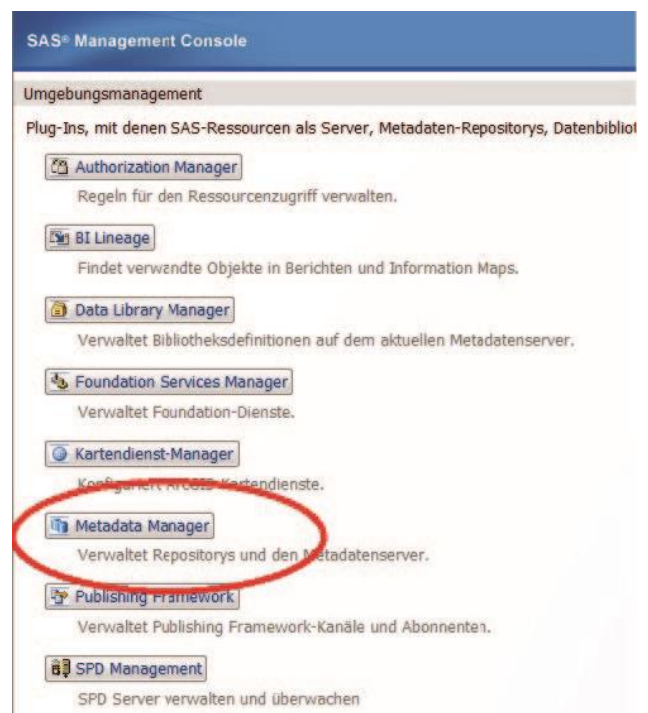


Abbildung 1: Metadata Manager

```
metaport=8561  
metauser='userid'  
metapass='password'  
metarepository='Foundation';
```

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

mayato GmbH
Am Borsigturm 9
13507 Berlin

www.mayato.com

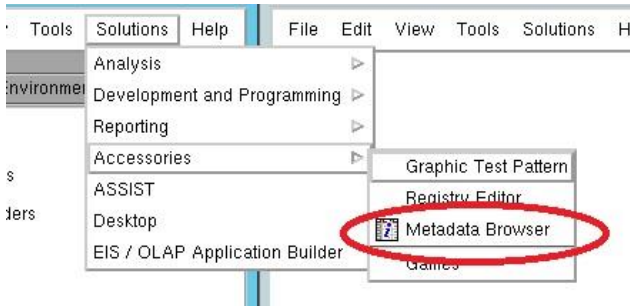


Abbildung 2: Metadata Browser

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

Name	Typ	Objekt-ID	Description	Modifiziert (GMT)
000_PROC_METALIB_Job		ASTUWP8.C0000005		18. Dezember 2014 21:48 Uhr
AB_achtstestprogramm (Job		ASTUWP8.C000001X	Generated by S.	25. Februar 2015 18:49 Uhr
AB_neuntestprogramm (Job		ASTUWP8.C000001Z	Generated by S.	25. Februar 2015 18:49 Uhr
Bspbaeckern_classcopy Job		ASTUWP8.C0000000		18. Dezember 2014 21:48 Uhr
DataCheck Job		ASTUWP8.C000000T		18. Dezember 2014 21:48 Uhr
DESTATIS_Bevoelkeru Job		ASTUWP8.C000000Z		05. Januar 2015 10:13 Uhr
DIFT Organization Dim Job		ASTUWP8.C000002D		19. März 2015 09:02 Uhr
DIFT Populate Order F Job		ASTUWP8.C000002C		19. März 2015 15:33 Uhr
DM_STUDIE_2015_BA Job		ASTUWP8.C0000007		18. Dezember 2014 21:48 Uhr
DQ_CARRIAGE_RETU Job		ASTUWP8.C0000027		10. März 2015 16:56 Uhr
ECC_DATABASE_V2 Job		ASTUWP8.C000000P		18. Dezember 2014 21:48 Uhr
ECC_DATABASE_V3 Job		ASTUWP8.C000000Q		18. Dezember 2014 21:48 Uhr
ECC_DATABASE Job		ASTUWP8.C0000000		18. Dezember 2014 21:48 Uhr
edrakt_testen Job		ASTUWP8.C0000002		08. 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 ein XML-Request.

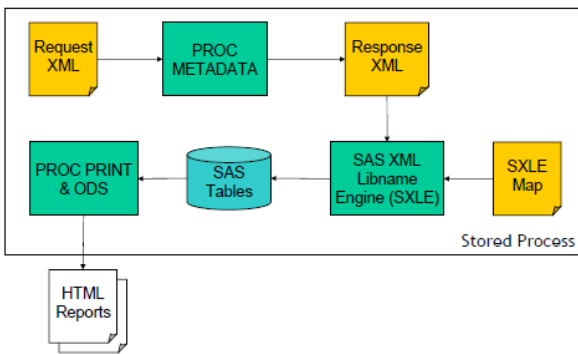


Abbildung 4: XML mit SAS®

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 weiter verarbeiten.

Ihr Ansprechpartner:

Melanie Hinz
melanie.hinz@mayato.com

mayato GmbH
Am Borsigturm 9
13507 Berlin

www.mayato.com



3 SAS® Data Steps

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

An dem Beispiel einer Jobliste mit Pfaden sieht man nun, dass die Objekte verschiedenen Typs zusammenhängen, sowie, dass wir ihre Eigenschaften auslesen können.

Man gibt den Namen des Repositorys an und initialisiert die Variablen.

```
OPTIONS metarepository="Foundation";
DATA Joblist;
  FORMAT name $64.
         uri_job $32.
         modified $32.
         uri_tree tree_name uri_par path
$255.
  ;
```

Mit einer Schleife über alle Objekte in der SAS®-Bibliothek, die vom Typ „Job“ sind, erhält man den URI jedes Jobs:

```
DO UNTIL (lrc lt 0);
  n+1;
```

Ihr Ansprechpartner:

Melanie Hinz
melanie.hinz@mayato.com

```
lrc =
metadata_getnobj("omsobj:Job?@Name ? '",
n, uri_job);
```

Zu jedem Job braucht man dessen Namen und wann er zuletzt geändert wurde:

```
IF lrc > 0 THEN DO;
  rc = metadata_getattr(uri_job,
'Name', name);
  rc =
metadata_getattr(uri_job,"MetadataUpdated"
,modified);
```

Jeder Job liegt in einem Ordner, ist also mit ihm assoziiert. Dieser hat natürlich ebenfalls einen Namen:

```
rc =
metadata_getnasn(uri_job,"Trees",1,uri_tre
e);
rc =
metadata_getattr(uri_tree,'Name',tree_name
);
```

Indem man jeweils den übergeordneten Ordner bestimmt, kann man den kompletten Pfad aufbauen.

```
urc =
metadata_getnasn(uri_tree,"ParentTree",1,u
ri_par);
path="/"||strip(tree_name);
do while (urc>0);
  uri_tree=uri_par;
  rc =
metadata_getattr(uri_tree,'Name',tree_name
);
  urc =
metadata_getnasn(uri_tree,"ParentTree",1,u
ri_par);
  path="/"||strip(tree_name)||path;
end;
OUTPUT;
END; * if lrc...;
END; * do until...;
KEEP name uri_job modified path;
run;
```

Führt man diesen Code aus, zum Beispiel in dem User Written Code-Knoten in einem DI-Studio-Job, sieht das Ergebnis wie folgt aus:

mayato GmbH
Am Borsigturm 9
13507 Berlin

www.mayato.com



The screenshot shows a workflow in DI-Studio with a 'Table Loader' job. Below it is a table titled 'View Data: T_JOBLIST (34 rows)'. The table has columns for job ID, name, modified date, job type, and path.

#	name	modified	jobtyp	path
5	best12	19Feb2015:14:08:51	Job	Project\mayato\Data Integration\020 Job\kompa
6	DQ_CARRIAGE_RETURN	10Mar2015:16:58:28	Job	Project\mayato\Data Integration\020 Job\DOCC
7	versprechen	18Mar2015:13:21:42	Job	Project\mayato\Data Integration\020 Job\kompa
8	Projects_memo_C3_Cas	09Mar2015:13:13:00	Project	
9	seditesprogramm (Generat...	03Mar2015:13:59:24	Job	Project\mayato\Data Integration\020 Job\kompa
10	seditesprogramm (Generat...	03Mar2015:13:51:49	Job	Project\mayato\Data Integration\020 Job\kompa
11	seditesprogramm (Gen...	05Mar2015:18:48:52	Job	Project\mayato\Data Integration\020 Job\kompa
12	AB_schleppprogramm (Gen...	05Mar2015:18:49:33	Job	Project\mayato\Data Integration\020 Job\kompa
13	impagation_job	17Mar2015:11:47:05	Job	Project\mayato\Data Integration\020 Job\kompa
14	besten_zustuehle	02Mar2015:16:57:59	Job	Project\mayato\Data Integration\020 Job\kompa
15	sed_audobieren	19Mar2015:16:29:17	Job	Project\mayato\Data Integration\020 Job\kompa
16	best11	12Mar2015:18:01:22	Job	Project\mayato\Data Integration\020 Job\kompa
17	best10	13Mar2015:17:15:44	Job	Project\mayato\Data Integration\020 Job\kompa
18	best9	13Mar2015:14:43:23	Job	Project\mayato\Data Integration\020 Job\kompa
19	best7	09Mar2015:18:26:36	Job	Project\mayato\Data Integration\020 Job\kompa
20	zusammenfassung_scheffe...	10Mar2015:14:22:58	Job	Project\mayato\Data Integration\020 Job\kompa
21	BESTATED_Bemerkung...	06Mar2015:18:13:41	Job	Project\mayato\Data Integration\020 Job\kompa
22	TI14ABC	18Dec2014:21:48:07	Job	Project\mayato
23	STAT_LINFAELLE_LAND_K...	18Dec2014:21:48:07	Job	Project\mayato\Data Integration\020 Job\Statist
24	STAT_BEVOEL_LAND_K...	18Dec2014:21:48:07	Job	Project\mayato>Data Integration\020 Job\Statist
25	STAT_BEVOEL_LAND_GE...	18Dec2014:21:48:07	Job	Project\mayato>Data Integration\020 Job\Statist
26	PROC_METADATA_306	18Dec2014:21:48:07	Job	Project\mayato>Data Integration\020 Job\Meta

Abbildung 2: DI-Studio-Job mit Jobliste

4 Weitere

Anwendungsmöglichkeiten

Den umgekehrten Weg kann man ebenso gehen. Man startet mit allen Ordnern und holt sich die Inhalte. Das können nicht nur Jobs sein, sondern auch andere

Objekte, wie Tabellen, Notizen, registrierte Dokumente, usw..

Des Weiteren gibt es Security Macros von SAS®, mit denen man Access Control Settings auswerten kann. Ein Beispiel ist, man lässt sich zu einem Ordner die Berechtigungen auf einem bestimmten Ordner anzeigen. Details dazu findet man auf der SAS® Homepage.

5 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

Ihr Ansprechpartner:

Melanie Hinz
melanie.hinz@mayato.com

mayato GmbH
Am Borsigturm 9
13507 Berlin

www.mayato.com



Kontaktieren Sie uns //

Ihr Ansprechpartner:

Melanie Hinz
melanie.hinz@mayato.com

mayato GmbH
Am Borsigturm 9
13507 Berlin

www.mayato.com