

Udskrevet er dokumentet ikke dokumentstyret.



## ETL-principper

Dokumentbrugere: <b>KONC-POEM-MT</b>	Dokumentansvarlig: <b>D-PØMQansvar</b>	Redaktør:  Version: <b>2.2</b>	Dokumentnummer: <b>12. 4.01</b>	Godkendt af: <b>JOEJB1308</b>  <b>22.09.2009</b>	Niveau:
---	---	---	------------------------------------	---	---------

- 1) Indledning
- 2) Generelle principper
  2. 1) Biblioteksstruktur
  2. 2) Flows
  2. 3) Processer
  2. 4) Transforms
- 3) Navngivning
  3. 1) Flows
  3. 2) Processer
  3. 3) Transforms
  3. 4) Excel
- 4) Dokumentation
  4. 1) Flows
  4. 2) Processer
  4. 3) Transforms

### 1) Indledning

I det efterfølgende vil principperne for opbygning af programmer i SAS Dataintegration Studio blive gennemgået (Herefter SAS-ETL).

Dokumentet anvendes i forbindelse med opbygningen af ETL programmer i udviklingsprocessen og ved efterfølgende review af programmerne.

Først gennemgås de generelle principper, herefter navngivningen og til sidst kravene til dokumentation.

### 2) Generelle principper

#### 2. 1) Biblioteksstruktur

Programmer og metadata for de enkelte modeller placeres i modellens bibliotek under '04 modeller'. I biblioteket for hver enkel model opbygges en mappestruktur, med følgende indhold:

- 01 Processer
- 02 MD manuelle data
- 03 TD temporære data
- 04 ABC grænseflade data
- 05 CD check data

Øvrige mapper i biblioteket defineres frit.

#### 2. 2) Flows

Flows trækker på en styrefil hvori det fremgår hvilke processer/jobs der skal køres i flowet.

Flows der vedrører MD rollen skal placeres i det overordnede 'MD flows' bibliotek.

#### 2. 3) Processer

ABC-processerne skal placeres under de enkelte modellers bibliotek i 04 modeller. Der oprettes et bibliotek '01 Processer'. Det anbefales at dette bibliotek underopdeles efter hvilken type output der dannes. F.eks.

- 0 Manuelle data
- 1 Costpooling
- 2 Hierarkier og accounts

- 3 Drivere
- 4 Rapportering
- 5 Andet

I øvrigt gælder følgende anbefalinger:

- Der er kun én primær outputtabel pr. proces (én ABC-tabel)
- Hvis processen bliver for kompliceret, deles den i to.
- Processer der ikke anvendes fjernes fra produktionsserveren.

## 2. 4) Transforms

Her gælder følgende:

- Der må ikke være løse ben, f.eks. to input hvor kun den ene er udfyldt.
- Der må ikke laves transformeringer i Loaderen. Kun 1:1 mapning.

## 3) Navngivning

### 3. 1) Flows

Flows navngives ved at angive et nummer først efterfulgt af et sigende navn. Nummereringen angiver rækkefølgen i forhold til de øvrige flows. Flows der kun vedrører en enkel delmodel skal have delmodel bogstav med i navngivningen.

F.eks.:

1\_flow\_manuelle\_data\_F

### 3. 2) Processer

En proces skal antage samme navn som output-tabellen. Ved flere output-tabeller skal processen navngives efter den primære. Processen der danner ABC\_F\_Dr\_manuelle skal f.eks. hedde ABC\_F\_Dr\_manuelle.

### 3. 3) Transforms

Ved navngivning af Transforms skal typen fremgå. Følgende princip gælder:

<type transform>\_<sigende navn>

F.eks. E\_udsøg.

De forskellige typer har følgende bogstav:

- E - Extract
- A - Append
- O - SAS sort
- S - SAS splitter
- J - SQL join

Loaderen navngives ikke da der ikke sker nogen behandling i denne transform.

DSB transforms navngives ikke da de allerede er navngivet.

### 3. 4) Excel

De manuelle data indlæses fra Excel. Da tabellen i SAS-ETL får navn efter navnet på fanebladet skal fanebladet i Excel hedde:

EX\_<modelbogstav>\_<type af manuelle data>

## 4) Dokumentation

Udover dokumentation i ETL som beskrevet herunder, skal komplicerede processer også være en del af modeldokumentationen. F.eks. beskrives dannelse af automatiske nøgler.

**4. 1) Flows**

Flows er dokumenteret i styrefiler. I 'egenskaber' til flowet beskrives kort hvad flowet indeholder.

**4. 2) Processer**

I 'egenskaber' til processen beskrives kort hvad processen gør.

**4. 3) Transforms**

I 'egenskaber' til de enkelte transforms beskrives hvad transformen gør.

**Revisioner**

Version	Godkendt	Revisions information
<a href="#">2</a>	27.05.2008	
<a href="#">2.1</a>	29.05.2008	
<a href="#">2.2</a>	22.09.2009	