

The HAWK has a good (SAS) view

Olivier LECONTE, SERVIER, Courbevoie, France

ABSTRACT

Like many pharmaceutical companies, SERVIER uses Clintrial™ for Clinical Data Management. In order to access the data in Clintrial™, SERVIER has chosen to use SAS® views.

This paper describes HAWK, an in-house application developed using SAS\AF, to create and manage these views.

The focus is on how Data Managers create views with high quality labels and formats using Clintrial™ metadata, and on the use of a batch tool to make daily updates of the views.

INTRODUCTION

Previously, the clinical data stored in Clintrial™ were validated using SAS. In 2002, we decided to access the clinical data by creating SAS views. This solution has the double advantage of being not only in real-time but it also guarantees data integrity. Today it is the only way our SAS users access clinical data.

The SAS® views had to have the same properties as the panels to which they were linked : same panel name, same item names, etc. We had to find a way to access the information on panel descriptions is stored in Clintrial™ metadata files.

The last step of this project was to identify a user-friendly interface to access the Clintrial™ metadata and enable Data Managers to initiate and create SAS views on their protocols. The end-result was an application we call Hawk.

USING HAWK

This section describes how Data managers use HAWK to set up their protocol.

DEFINE A STUDY

The first step is for the user to enter the protocol to be accessed.

Options de génération :

Etude : [dropdown]

Instance Clintrial : [dropdown]

☐ Etude active

☐ Mise à jour automatique

☐ Génération des vues STATS

☐ Générer les extractions

Fréquence : [dropdown] (en jours)

☐ Présence de données Meddra et Whodil

☐ Lien automatique entre les tables d'AE

☐ Générer la table ADVEN_SMQ

☐ Générer les tables ADVEN_MUL et ADVEN_SCC

Exclude Panel

Répertoire de stockage (si différent du numéro de protocole) : [text field]

Interlocuteurs :

[text field] [right arrow] [left arrow]

Mails: [text area]

Note: la forme du mail doit être AXTR_CB, il faut rentrer au moins un correspondant.

[floppy disk icon] [Quitter]

After the protocol has been entered, the user can specify options to :

- Request daily monitoring of the database structure and status to know if the views/extracts have to be regenerated
- Generate specifics views for the statisticians without Coding or items internal to Clintrial™
- Create specifics views for Adverse Events if the statisticians want to perform SMQ, SSC, or multi-axial MedDRA analyses.
- Generate extraction on a daily, weekly or monthly basis.

For each protocol, the data-manager defines the contact(s) who will receive an e-mail each time the view is updated. In this way, SAS programmers are automatically warned that the database has evolved, which might have an impact on their programs.

EXTRACTING DATA

The screenshot shows the HAWK web interface. At the top, there is a blue header with the 'SERVIER' logo on the left and the 'HAWK' logo in the center, accompanied by a falcon icon on the right. Below the header, there is a dropdown menu labeled 'Etude :'. Underneath, there is a text input field labeled 'Chemin :'. A central panel contains options for 'Type d'export' (radio buttons for 'csv', 'txt', and 'dataset') and a list of checkboxes for various data handling options: 'Données codées (Ex: YES,NO)', 'Données décodées (EX: 1,0)', 'Avec Meddra/Whodc', 'Sans Meddra/Whodc', and 'Exclure les items system'. An 'Exclure Panel' button is located below the 'Type d'export' options. At the bottom of the interface, there are two buttons: 'Extraire' (with a document icon) and 'Quitter' (with a right arrow icon).

HAWK allows extraction of data in SAS datasets and Excel formats. Panels can be excluded from these extracts.

If the data are sent to a sub-contractor that does not have licences for MedDRA or WHO-DRL, HAWK automatically transforms the MedDRA and WHO-DRL numerical codes into terms. The corresponding formats are also deleted in the format catalog.

AUTOMATIC UPDATE

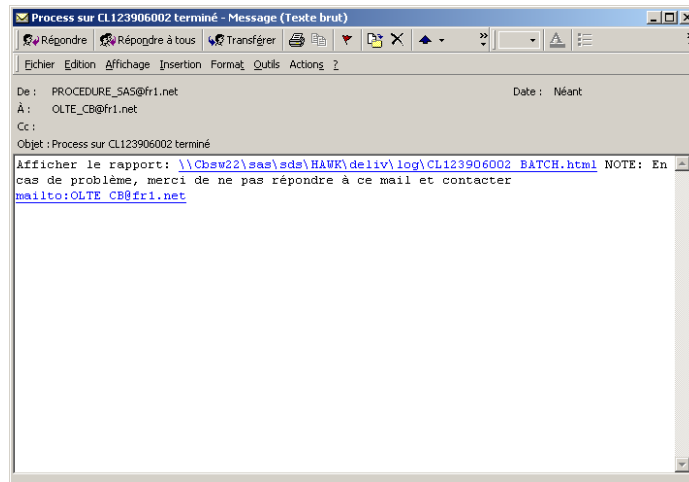
Two batch jobs are submitted on a daily basis. The first, called BATCHNIGHT, is used to identify locked studies and to extract them automatically. The second, called BATCHAM, is executed every morning, and updates protocol views if needed (because of a new panel, new item, etc.). It also performs SAS extractions if requested by the user.

A third batch job, called REALTIME, is executed every ten minutes and is used to create views and datasets on demand.

These batches are submitted using Control M software. Before submission, Control M verifies that all the servers (Clintrial + SAS) are available. If not, the batches are re-submitted ten minutes later.

The use of Control M allows views and datasets to be created for a specific user account, managed by the IT teams, and to stored the views and datasets in secured and protected folders.

Each time a study is updated, the contacts receive an e-mail informing them that there has been a change.



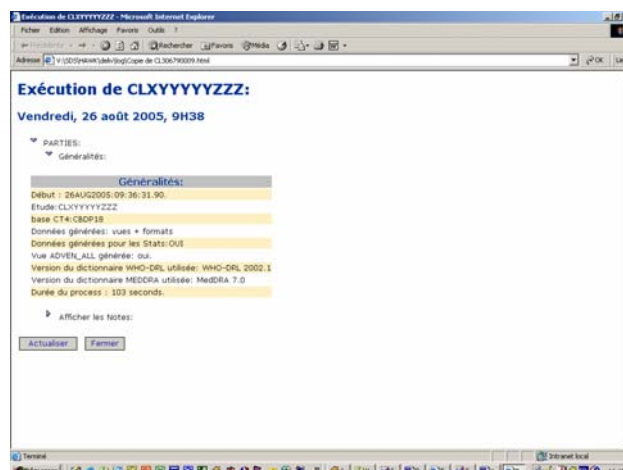
This mail is sent using the SOCKET FILENAME which is linked to the SMTP server :

```
filename smtp socket "CB-SMTP.FR1.GRS.NET:25";

data _NULL_;
  file smtp;
  x = sleep(5, 0.1);
  put 'ehlo localhost' '0D'x;
  x = sleep(5, 0.1);
  put "mail from: &from" '0D'x;
  x = sleep(5, 0.1);
  put "rcpt to: &to" '0D'x;
  x = sleep(5, 0.1);
  put 'data' '0D'x;
  x = sleep(5, 0.1);
  put 'Mime-Version: 1.0' '0D'x;
  put 'Content-Type: multipart/mixed;boundary="===== _-
1366791890===== "' '0D'x;
  put "To: &to" '0D'x;
  put "From: &from" '0D'x;
  put "Subject: &subject" '0D'x;
  put '0D'x;
  put '----- _-1366791890===== ' '0D'x;
  put 'Content-Type: text/plain;' '0D'x;
  put "Content-Disposition: inline" '0D'x;
  put '0D'x;
  put "&body" '0D'x;
  %if "&file" ^= "" %then %do;
    put '----- _-1366791890===== ' '0D'x;
    put 'Content-Type: text/plain;' '0D'x;
    put "Content-Disposition: attachment; filename="&filename"" '0D'x;
    put '0D'x;
    do until(slutt);
      input;
      put _infile_;
    end;
  %end;
  put '----- _-1366791890===== ' '0D'x;
  put '.' '0D'x;
  put '0D'x;
  x = sleep(5, 0.1);
  put 'quit' '0D'x;
  stop;

run;
```

Using this mail, the users can access the HTML log which described the protocol, its options, and the reason for the update :



USING CLINTRIAL METADATA

GENERAL POINTS

All the characteristics of a protocol (the panels, the items, etc.) are stored in Clintrial™ Metadata protocols called CTS and CTSD. We developed a macro-procedure capable of requesting the characteristics of a Clintrial™ protocol and creating views using the information in the protocols.

The list of views to be created is requested from CTS.PANEL metadata panel.

Using the CTSD.ITEM metadata panel, we are able to get all of the characteristics of an item (name, sasname, order, type and the codelist associated.)

Most of the time, the item name and item sasname are the same. With version 8, we can use an item name with more than 8 characters (up to 32).

Using the panel information, we can write an SQL procedure to create the view :

```
proc sql;
create view viewd.pasta as
select CT_RECID as CT_RECID label='RECORD IDENTIFICATION',
       ENTRY_DATETIME as ENTRY_DA label='ENTRY DATETIME' format=DATETIME20,
       STATUS as STATUS label='STATUS OF RECORD',
       ENTRY_ID as ENTRY_ID label=' USER ID',
       MERGE_DATETIME as MERGE_DA label=' LAST MODIFICATION DATE' format=DATETIME20,
       TRIM(STUD_ID) as STUD_ID label='STUDY' format=$3 informat=$3,
       TRIM(COUN_ID) as COUN_ID label='COUNTRY' format=$ISOCODS. informat=$3,
       TRIM(CENT_ID) as CENT_ID label='CENTRE' format=$4 informat=$4,
       TRIM(PATID_ID) as PATID_ID label='PATIENT IDENTIFICATION' format=$5
       GOTRT as GOTRT label='STUDY TREATMENT GOING ON' format=YESNOS. informat=1.0,
from CPROT.PASTA_UPDATE
order by CENT_ID asc, PATIDENT asc, VISIT asc;
quit;
```

The context items, such as CT_RECID, are not described in the metadata panel. They are automatically added to the views.

The sort order is defined by HAWK depending the name of the panel and the items it contains.

MANAGEMENT OF NUMERIC FORMATS

For numeric items such as laboratory parameters, the format descriptions in SAS and Clintrial™ are not equivalent.

A float item in Clintrial with a format 4.2 has 4 digits before the comma and 2 after. The corresponding format in SAS is 6.2. The negative sign is not counted as a character by Clintrial™ but it is in SAS.

So the final formula to determine the SAS format is : if Clintrial™ format is **M.N** then SAS format is **M+N+1.N**.

MANAGEMENT OF DATES

Clintrial™ can stored only datetime date. We use the item sasname to indicate if the data is a date, a time or a datetime. If the item sasname ends with DTI, it concerns a datetime. If the item sasname end with DAT, it concerns a date. The DATEPART function is used to extract the date. If the item sasname ends with TIM, it concerns a time. The TIMEPART function is used to extract the time.

In order to avoid a warning message from the DATEPART or TIMEPART functions when the date value is null, we use the following code :

```
case TESTDAT is null
  when 1 then .
  else DATEPART(TESTDAT)
end as TESTDAT label='TEST DATE'
```

For a each date, we add a character date which allows an incomplete date to be identified, using the fact that day, month, and year are stored in specific variables and by applying the following algorithm :

```
case (((2**0)*(YVIS not = .))+((2**1)*(MVIS not = .))+((2**2)*(DVIS not = .))
  when 0 then ''
  when 1 then '....'||PUT(YVIS, Z4.)
  when 2 then '...'
  when 3 then '..'||UPCASE(PUT(MDY(MVIS, 1, 2000), MONNAME3.))||PUT(YVIS,Z4.)
  when 4 then '...'
  when 5 then PUT(DVIS, Z2.)||'..'||PUT(YVIS, Z4.)
  when 6 then '...'
  when 7 then PUT(DVIS,Z2.)||UPCASE(PUT(MDY(MVIS,1,2000),MONNAME3.))||PUT(YVIS,Z4.)
  else ''
end as VISDATC label='VISIT DATE'
```

In this way, SAS users can easily detect an incomplete date in the data listings.

FORMATS

Formats are generated using Codelists metadata. This allows the same decode for data under Clintrial™ and SAS®.

There are 3 steps :

- Requesting name, value and code of codelist from metadata panel CTS.AGGREGATED_CODES
- Requesting type of codelist from metadata panel CTS.CODE_INDEX
- Merging the 2 datasets and input of the resulting dataset in a PROC FORMAT CNTLIN= syntax.

Medical dictionary formats such as MedDRA or WHO-DRL are pre-generated for each version. We have an in-house metadata panel which references the versions used for each protocol. Using this in-house metadata panel, HAWK copies the selected version in the protocol formats catalog.

METADATA BACKUP

For each protocol, a copy of its metadata is stored in SAS datasets in order to be able to detect changes. This metadata is also used by the other standard SAS programs each time information concerning the data structure of a panel is required. It allows us to save time by not performing PROC CONTENTS or a request on the SASHELP.VCOLUMNS table.

OTHER DATA

HAWK generates views to Monitor and Numerized data that can be used in SAS® programs to track CRFs and DCFs.

CONCLUSION

Thanks to HAWK, our SAS users have now high-quality views that are automatically updated.

We plan to improve the Export functionality in the next few months and to add a new module in order to manage Analysis datasets. With SAS® version 9, we may be able to provide CDISC compliant extracts.

REFERENCES

Leconte, Olivier, *Do you need second sight to use the LIBNAME ORACLE*, Proceeding of PHUSE 01, paper DM06, 2005.

CONTACT INFORMATION (HEADER 1)

Your comments and questions are valued and encouraged. Contact the author at:

Olivier LECONTE
Institut de Recherches Internationales SERVIER
8 Places des Pléiades
92415 Courbevoie Cedex
Work Phone: 33 1 55 72 72 08
Email: olivier.leconte@fr.netgrs.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.