

Paper 27-25

SynCoder: A SAS/AF® Application for Implementing MedDRA Coding in Clinical Trials

Deborah Testa, Seven of Nine Systems, Studio City, CA

Russell Holmes, Synteract, Inc., Encinitas, CA

ABSTRACT

This paper presents SynCoder, a Version 7/8 SAS/AF® application that implements the new MedDRA international medical terminology for coding adverse events in clinical trials. The presentation will be accompanied by a brief demonstration of three of the application's modules: batch coding using a synonyms data set and a data mart of MedDRA hierarchical terms, an efficient search engine for browsing the hierarchies, and on-line coding of terms requiring human decision making. The application demonstrates the superiority of new features (both in the user interface and in the component technology) in SAS/AF software in Version 7 and 8 over earlier versions.

INTRODUCTION

In June of 1999, the publication of a new international medical terminology – highly structured and more intricately hierarchical than most current coding schemas used in drug development – resulted in a major impact on data management systems and procedures in pharmaceutical and biotechnology companies.

The new terminology, called MedDRA for short, was the result of four years of work by an international committee working under the auspices of the International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use. That committee was tasked with developing a single internationally accepted medical terminology (a coding system) for regulatory purposes. As such, the terminology replaces familiar coding schemas such as COSTART, WHO-ART and ICD-9 with which professionals in drug development have communicated for years, and around which companies had built proprietary software systems and long-standing procedures.

Any time a profession undergoes a major change in basic terminology such as that posed by MedDRA, software systems must change to accommodate.

Synteract, Inc., a contract research organization providing comprehensive clinical development services, responded to the publication of the new MedDRA terminology by developing a SAS/AF application to provide the computing environment for implementing the MedDRA codes in the context of the clinical trials it was managing for its clients. The result is SynCoder, as SAF/AF Version 7/8 application that fully implements MedDRA.

SYNCODER DATA STRUCTURES

The foundation of this application is a data mart of SAS data sets and MDDB's, into which SynCoder loads the 14 ASCII files distributed to subscribers by the MedDRA service organization. SynCoder's data structures denormalize the data and optimize it both for fast retrieval (using indexes) and for clinically meaningful display (using the features of FRAME entries). Our strategy was to offload as much of the data processing as possible to the process that creates this data mart – even at the cost of a great deal of disk space – in order to minimize the work required during

human use of the system. The system minimizes the human effort involved in coding wherever possible. Thus the data mart stores not only the text as it was published by the committee (upper and lower case) but also an upper case variant as well to facilitate indexed reads. SAS does a more resource-intensive sequential read if a WHERE-Expression contains a function other than TRIM or SUBSTR, so redundant storage facilitates fast retrieval without intruding on the style of display chosen by the ICH committee.

SynCoder's data library also contains an MDDB created by SAS/MDDB Server™ for each supported terminology. Since full implementation of MedDRA is not required until 2001, SynCoder implements COSTART as well as multiple versions of MedDRA. The MDDB's support on-screen drilling through the hierarchies.

Discovering the optimum structure for storing MedDRA terminology in a SAS data mart customized for the SynCoder application was the primary challenge -- and the primary feature – for the product, and we continually tweak our decisions as new requirements surface.

SEARCH THE DICTIONARIES

Access to the data marts is provided by SAS/AF FRAME entries that organize several different modules.

First, an on-line search engine provides users with the ability to search – given a particular medical term -- for the term's preferred term and paths in the hierarchy of the selected coding terminology. The MedDRA terminology contains cross-references to COSTART terms, and these can be accessed on screen via SynCoder.

Because terms frequently appear in more than one hierarchy, the application displays each relevant path in the hierarchy with the primary one distinguished by the color in which the text appears on the screen. The hierarchical nature of the terminology made the organizational chart object a natural viewer object for the search results. And because the underlying data structure has been optimized, results return from a data mart with nearly half a million observations in about one second.

A second way to access the dictionaries is to use the browse facility to drill through a multi-dimensional database that has been produced from the data mart detail tables. This method of access is useful for browsing from the top of the hierarchy down, whereas the search engine is useful for extracting lower-level hierarchies based on search parameters.

Via SAS/Intrnet™, Synteract will provide its clients with extranet access to this search engine as a service in the next release of the software.

AUTOCODER

Second, SynCoder's AutoCoder module performs batch coding of terms for which there is a simple match in MedDRA or COSTART terminology. Because AutoCoder first accesses a proprietary

data set of synonyms storing lexical and spelling variants accumulated over time, the chances of finding a match increase over time. The AutoCoder takes the term as it was reported verbatim by a clinical investigator, matching it first against this synonym data set to reduce variation, then matching the result of that operation against the MedDRA data mart.

Because of SynCoder's heavy use of indexes, the AutoCoder process is extremely quick. A log data set maintains a record of each batch process for auditing purposes.

ONLINECODER

An autocoder will never code everything – human decision-making is thankfully still both an asset and a requirement. SynCoder provides the computerized mechanism for facilitating and tracking such decision-making.

In the case of terms for which AutoCoder could not find a simple match, then a trained clinical data specialist must decide between multiple legitimate choices. SynCoder presents uncoded observations next to those legitimate choices -- with the full MedDRA hierarchy represented in the same style as the search engine – and lets the user register a selection with a mouse click. The OnLineCoder module responds to that mouse click by populating empty variables in the observation on screen with the data selected.

As each observation is coded, SynCoder accumulates a list of variant term/standard term pairs, and at the end of the session the application offers the user the opportunity to add these pairs to the synonyms data set.

Like the AutoCoder module, OnLineCoder maintains a log of sessions.

CONCLUSION

SAS Version 7 and 8 provided several applications development features that were key to the success of SynCoder, including the coding ease provided by dot notation, the lightning-fast results of attribute linking, the new SAS/AF development environment, and the attractive nature appearance of new controls. Because the application implements model/viewer concepts, much of the coding is far more generic and efficient than applications we had previously developed in earlier version. The interface in particular, with its native Windows look-and-feel, gets extremely positive response from users. We think the application demonstrates the clearly dramatic improvement that has occurred in SAS/AF since Version 6.

REFERENCES

For SAS/AF developers making the transition from Version 6 will find the following resources helpful:

1. [SAS® Guide to Applications Development](#) (1999: SAS Institute, Inc., Cary, NC). This new publication is an excellent guide to Version 7 features of SAS/AF.
2. Wilson, Steven A., "An Application Developer's Roadmap for Moving to Version 7 in SAS/AF," *Proceedings of the Twenty-Fourth Annual SAS® Users Group International Conference*, 20, 119-124.

ACKNOWLEDGMENTS

SAS, SAS/AF and SAS/Intrnet are registered trademarks of SAS Institute Inc. in the USA and other countries.

CONTACT INFORMATION



Deborah Testa, Ph.D.
Seven of Nine Systems
Studio City, California
(818) 766-0186
dtesta@sevenofninesystems.com
www.sevenofninesystems.com



Russell Holmes
Vice President
Synteract, Inc.
Encinitas, CA
(760) 634-1002
rholmes@synteract.com
www.synteract.com