

# **TrialDB: A Web-based Clinical Study Data Management System**

## **AMIA 2003 Open Source Expo**

CA Brandt M.D., M.P.H., AM Deshpande, M.D., C Lu, M.S., G Ananth, M.S.,  
K Sun M.S., R Gadagkar M.S., R Morse<sup>2</sup>, C.Rodriguez, M.D.

PL Miller M.D.,PhD, PM Nadkarni M.D.

Center for Medical Informatics

Yale University School of Medicine, New Haven, CT  
Dept of Biostatistics, Massachusetts General Hospital<sup>2</sup>

### **INTRODUCTION**

Clinical Study Data Management Systems (CSDMSs) are a class of software that support centralized management of data generated during the conduct of clinical studies. Commercial CSDMSs include Oracle Clinical, ClinTrial and MetaTrial. Such systems, which are typically deployed at an institutional or organizational level, must accommodate diverse types of data from different clinical domains that is generated by different groups of clinical investigators. Large-scale CSDMSs typically employ a high-end database engine that is usually accessed over an intranet or the Internet using Web-based technologies. CSDMSs in institution-wide use for a variety of clinical domains are best served by entity-attribute-value (EAV) modeling for the clinical data: all the commercial CSDMSs that we are aware of use EAV design. However, de novo development of EAV databases for data management is a challenging task. A large body of generic metadata-driven code must be developed before a basic EAV application can be written. Clearly, the availability of pre-existing software with the requisite functionality would be very valuable. We will discuss the benefits of such software being in open-source form.

### **BACKGROUND**

The development of TrialDB in 1997 was necessitated by the limited functionality of commercial CSDMS software at the time. Most of the original problems with commercial software still remain:

High Per-Seat Costs: These are typically on the order of \$15-20 K.

Proprietary Architecture: This is typically combined with limited or non-existent developer-oriented documentation. Many of the costs comes from consultancy services to customize the software to the needs of particular clients.

Limited Built-in Functionality: Current commercial CSDMSs support data entry into the CSDMS through the keyboard: streamlined bulk data import and export is rarely built-in.

Limited or No Support for Binary Data: Clinical studies are increasingly likely to store Binary Large Object (BLOB) data such as images.

Inability to Customize the Software for Special Purposes: One feature that was specifically needed in TrialDB for the Cancer Genetics Network was the ability to display interactive pedigree diagrams over the Web, where clicking on a subject's icon would bring up the subject's data. This is niche functionality that commercial CSDMS vendors will not provide because there is not enough demand to justify the outlay in R&D.

### **CURRENT STATUS**

TrialDB is a fully functional CSDMS that is in operational use at multiple sites. It continues to evolve to meet the requirements of diverse active users. The documentation and code are freely available to investigators in academia at <ftp://custard.med.yale.edu>.

### **References**

1. Nadkarni PM, Brandt C, Frawley S, et al. Managing attribute-value clinical trials data using the ACT/DB client-server database system. Journal of the American Medical Informatics Association 1998;5(2):139-151.
2. Nadkarni P, Brandt C, Marengo L. WebEAV:Automatic Metadata-Driven generation of Web Interfaces to Entity-Attribute-Value Databases. JAMIA 2000.
3. Brandt CA, Nadkarni P, Marengo L, Karras BT, Lu C, Schacter L, Fisk JM, Miller PL. Re-engineering a Database for Clinical Trials Management: Lessons for System Architects. Controlled Clinical Trials 21:440-461,2000.