
DataFax integration with relational database systems: Reporting and alternative data entry options

Mark Lowerison, Christina Ma
Clinical Research Unit,
Hotchkiss Brain Institute,
University of Calgary,
Calgary Alberta

CRU CLINICALRESEARCHUNIT



About the Clinical Research Unit

- The Clinical Research Unit (CRU), is a core facility at the Hotchkiss Brain Institute
- The CRU is designed to assist investigators developing and launching studies and/or clinical trials
- Special emphasis is placed on supporting the translation of research knowledge into practical treatment tools to benefit the health of the general population

About the Clinical Research Unit

- Established in 2008 we are still spending time on process and policy development
- We are still refining our services and policies as we work towards our first planned audit in mid 2011
- CRU audit and review goals are being supported by technologies and protocols outlined in today's talk

CRU Data Collection/Reporting Overview

- Two branches for our data collection services
- DataFax for all hypothesis driven data acquisition and storage
- Web applications for some non-hypothesis driven data collection objectives
- Web based reporting for both DataFax and Web databases

Overview of our DataFax Environment

- Currently using 3.8.3 in planning stages of an upgrade to latest version.
- Currently running on Solaris but likely to use Linux with future implementations.
- Primary driver in this decision is Virtualization Environment at U of C
- Excited for new features of 4.X versions of DataFax

Overview of our Web Application Environment

- MySQL datastore
- AJAX Java/JavaScript user interface
- Entirely virtualized hosting environment with UI and DB services being hosted independently
- Capacity for development of multi-tenant and multi-role data collection tools
- All within the secure computing and storage areas of the U of C HiiTec Knowledge Hub

Overview of Reporting Environment

- Rely heavily on BIRT (Business Intelligence and Reporting Toolset)
- Web Applications generally contain report generation tools as part of the main user interface
- DataFax Databases currently only use BIRT for Audit and QC reports however we are starting development of BIRT based reporting for standard instruments in our library

Hooking it all together

- In order to facilitate web based reporting or integration of DataFax OCR into our web applications we need to achieve a few goals
 - Automate Dumping of DF data to RDBMS
 - Develop tools to migrate DF structured data to Web Databases (in a general context)
 - Outline and alleviate security considerations, in particular consideration of privacy guidelines

Automating DFSQL Load

- Dynamic Dump, Studies List comes from DFstudies.db
- Additional tables created by parsing DFschema allowing storage of plate numbers and descriptions and variable names and descriptions
- This info is useful for labeling variable names in statistical packages and this approach allows re-usable code to produce these labels

Migrating data from DataFax to Web Applications

- We want an easily re-usable solution with minimal manual field matching
 - Solution: Two approaches.
 - Start with specific scripts to manage migration of standard forms (Standardized demographics, frequently utilized instruments etc.)
 - Work towards a more generalized approach using MySQL's "describe" along with standardized table formats (particularly variable naming conventions) to facilitate generalized transformation

Outline and Alleviate Security Concerns

- Privacy of patient identifiable information is paramount
- When developing tools for use in clinical settings running strictly anonymized data is not feasible
- A variety of technologies protect privacy:
 - SSL Encryption Two-Token Authentication
 - Firewalls Domain Locking Password Ageing
 - Password Strength Monitoring SSH Pairing
 - These are major considerations and the rate limiting factor in most of our development timelines

Where are we...

- We now have automated schedulable backups to relational database format with no human intervention for all of our DataFax databases
- Information is being automatically parsed from DFStudies.db and DFSchema.db text files and being stored in database

Where are we...

- All our web app data lives in the same format with similar guidelines for record validation, etc. With the major omission of QA and Audit tools
- We have one project (A Surgical Registry) that is running as a hybrid DataFax-Web Application

Where next...

- Expansion of reporting for standardized instruments (reusable)
- Generalization of DataFax to Web database migration functions
- Investigation into presentation of raster and DFStatus components in Web Applications
- Expansion of QA and Audit monitoring tools into standalone global studies management reporting application

Where next...

- Delve into world of audit and quality control of our web applications taking queues from DataFax
- Development of a set of statistical programming language tools to take advantage of this data consolidation

Concluding Thoughts

- We have a lot of work left to do but we are happy to find that time spent on these types of tasks ends up benefiting the unit in more than one way:
 - Data Security and Redundancy
 - Data Accessibility without need for iDF or staff involvement (very sparingly used)
 - Facilitation of near real time metrics
 - Development of new products to service more diverse needs of clients

Questions

- **Contact info:**

Mark Lowerison, Statistician/Software Specialist
Clinical Research Unit, Hotchkiss Brain Institute, University of Calgary
mwloweri@ucalgary.ca
403.210.6375

Christina Ma, Manager
Clinical Research Unit, Hotchkiss Brain Institute, University of Calgary
chrma@ucalgary.ca
403.210.6372

<http://clinicalresearchunit.org>

cru CLINICALRESEARCHUNIT