

---

## An Integrated Approach for DataFax Federation and Collaboration

Jon Vandegrift, Hanming Tu  
Premier Research, A Division Of CRC Development, Ltd.  
30 South 17th Street, Philadelphia, PA 19103  
<http://www.premier-research.com/>

### Content

---

- Problem in Reporting
- Solution: Integration
- DataFax Federation
- DataFax Collaboration
- Conclusions

## The Problem

---

- Reporting for multiple phase and multi-system trials is a difficult task.
- Data collected in DataFax system may have to be in different studies.
- Some sites may have elected to use EDC system for data collection.
- Some data may exist in project management system.
- How can we put all the data into one report, especially for our Best Pharmaceuticals for Children Act - Coordinating Center (BPCA-CC, <http://www.bpca-cc.com/>)?

## Traditional Approach - Data Sets / Data Marts

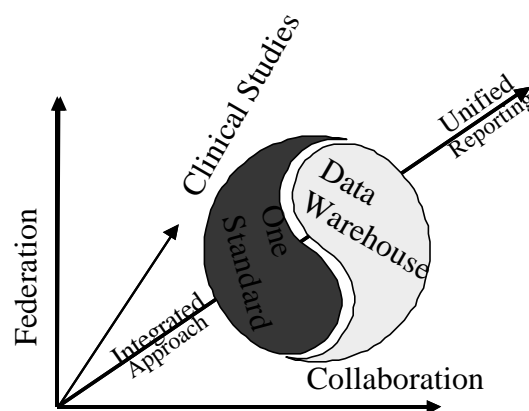
---

- **SAS Approach:**
  - Run DFsas to get data out of DataFax
  - Import data sets to SAS from other systems such as EDC, PPM, Lab, etc.
  - Crossing checking and merging in SAS
  - Generate reports, tables and listings in SAS.
- **Oracle Approach:**
  - ETL data from various systems including DataFax into Oracle, one study per data mart
  - Report from the single data mart
- **Limitation: no common standard, difficult conducting cross study reporting.**

## Integrated Approach - Data Warehouse

- Adopt a common industrial standard – CDISC.
  - Form a standard committee
  - Re-design all CRFs: use the standard variable and convention suggested by CDISC model
  - Require our EDC partner to use the same standard
- Build a centralized clinical data repository – data warehouse.
- Collaboration: Develop ETL to integrate with EDC, PPM, Oracle clinical, etc.
- Federation: Develop tools to generate reports from the data warehouse.

## Visualization of Integrated Solution



## Build a Data Warehouse

---

- Create schema based on CDISC domains in Oracle
- Build meta-database (code list) for all the controlled variables
- Build hierarchy for OLAP drill-down paths

## Collaboration Path

---

- Three stages for collaboration:
  - Data Quality: systems to systems
  - Information: people to systems
  - Knowledge: people to people
- DataFax collaboration is the very first step in the collaboration path – try to get quality data from various systems.

## DataFax Collaboration

---

- No APIs exist in DataFax, so it is not possible for direct collaboration.
- Indirect collaboration through the standard and data warehouse
- ETL all the plates from DataFax into data warehouse
- ETL all the CRFs from EDC into data warehouse
- ETL all the data from PPM into data warehouse
- ETL all the data from other systems if needed
- Will DataFax version 3.7 be helpful in this?

## Collaboration Example - LifeTree EDC

---

- We have LifeTree as our EDC partner in BPCA-CC project.
- We use LifeTree's ICTMS to collect data on the web.
- LifeTree places exported data sets in CDISC format on their ftp server.
- We transfer the data sets through a secure FTP and load them into Oracle database.

## Federation Path

---

- Three levels of federation
  - Structure: tables and columns
  - Code list: the same value means the same cross studies for controlled variable
  - Tools: Get the same presentation for the same report from multiple studies.
- DataFax federation is highly dependent on the database structure and standards embedded in it. We are in the process of acquiring Oracle OLAP tool.
- We just share a tool that we develop while we did not have the concept of adopting a standard and building a central data warehouse.

## Federation Example - PDF Generation

---

- In our registry studies, our clients usually want to see reports from multiple studies and need to see CRFs cross studies in one PDF file.
- We enhanced our PDF generation tool DF\_ATpdf to include the capability to access images from multiple studies.
- This is direct federation at tool level from DataFax.

## PDF Generation: UI and Flag

DataFax Web Reporter

Report source:	<input checked="" type="radio"/> DataFax Generic <input type="radio"/> Study Specific	<input checked="" type="radio"/> Local <input type="radio"/> Delayed <input type="radio"/> Real time <input type="checkbox"/> Verbose
Select studies:	001 - Mallin 1092-96-01	<input checked="" type="checkbox"/> Jazweb output <input type="checkbox"/> View Codes <input type="checkbox"/> Save
Select a report:	DF_ATpdf - : Create PDF with bookmarks (formerly DF_runDFpdf)	
Set options:	-s <study> ...the study number	
Actions:	-S primary -D P -a d -s <study>	
	<input type="button" value="Run Report"/>	<input type="button" value="Exploit"/> <input type="button" value="Refresh"/> <input type="button" value="Reset"/>

-s study numbers, required. For study on a remote server, you can specify as the following: SN:SVR:DataFax\_DIR.  
 To add a secondary study, use this input format:  
 SN:p\_SVR:p\_DFDIR,s\_SN:s\_SVR:s\_DFDIR  
 For example, to combine study 82 and a remote study 83 into PDF creation using '-s 82,83:df-svr2:/opt/datafax'

## PDF Generation: Example

- DF\_ATpdf has the following features:
  - Integrated into DataFax report: run within DataFax, in command line, and in web reporter.
  - Inherent all fine selection from DFexport.rpc
  - Divide PDF files into a CD full and linked all the files with a HML index file
  - Transfer PDF files to a burn station
- See example [here](#)

## Conclusion

---

- It is a long journey to achieve the two goals of integration: DataFax collaboration and Federation.
- The first step to start the journey is to adopt a common standard (soft infrastructure).
- The second step is to build your data repository – data warehouse (solid infrastructure).
- The two infrastructures pave a wider road for building an integrated solution for clinical data management.