

---

# DataFax 4

Clinical DataFax Systems Inc.

---

## Overview

- **Why?**
  - Motivation for DataFax 4
- **Implementation Details and Demonstrations**
  - Databases
  - Investigative Sites
  - Workflow
  - CRF Review
- **Changes since DFUG 2001 presentation**
- **Progress Report and Schedule**
  - Issues that need feedback
  - Licensing
  - Alpha and Beta Use

## Motivation for DataFax 4

---

- MS Windows support
- Standard database (Oracle, SQL Server) support
- Simplify modification of in-progress study database
- More key fields
- More than one QC note per field
- User extensibility in coding
- User roles
- Fine-grain permissions

## New Requirements

---

- Wide Area Network use
- Web-based publishing of data / reporting
- Audit why a data change was made – 21 CFR Part 11
- Heck, audit everything
- QC notes that reference multiple fields
- More meta data

## CDSI Requirements

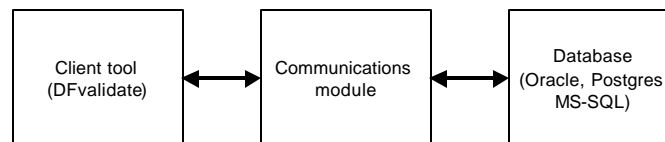
---

- Ongoing support for DataFax 3
- Simultaneous development on Windows, Solaris, HP-UX, and Linux platforms
- Hide differences between operating systems and databases from end user
- Leverage capabilities of database products
- Enforce permissions at the database level
  - Prevents subversion by a third-party application
- Allow ODBC access to database

## Architecture

---

- Three components: client tool, communications module and database backend
- More efficient (faster) WAN performance
- Encryption can be built-in
- Smaller client code, no "DLL Hell"




## Database Structure

---

- **Normalization**
  - Data modules (e.g. medications, AEs, etc.)
  - Plates are composed of 1+ data modules
- **New Keys**
  - Visit Repetition: repeat and interim assessments
  - Plate Repetition: continuation pages at the same visit
- **New Data Types**
  - Time, floating point, 3 of 9 bar code
- **Inter-study access is possible**
  - Copy data/meta data from a centralized library

## Normalization

---



DataFax #254
Plate #003
Sequence #001

Patient Number	99002	Patient Initials	JAM
----------------	-------	------------------	-----

**MEDICAL HISTORY**

	No		Yes		→	Duration <small>Years &amp; Months</small>		Currently on Treatment		Resolved and/or Controlled		
	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
1. Hypertension	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	→	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
2. Atrial Fibrillation	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	→	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>
3. Other Cardiac EG Arrhythmia	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	→	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. Valvular Heart Disease	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	→	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

## Normalization

---

- Modules contain the data fields and may be referenced on multiple plates
- A plate may contain multiple instances of the same module (for repeating data)
- Each module is associated with an SQL table, the fields with the columns of the table
- Module does not restrict field layout
- Permissions can be associated with module, not just plate

## Meta Data

---

- Everything is stored in the database
- In addition to a data value, missing code, or partial date, each field may have:
  - investigator comment (text field)
  - reason for value entered in the database (text field)
- Edit checks, visit scheduling, page map, etc.
- Table for defining user properties and storing those properties for:
  - study
  - site/subject
  - visit
  - plate
  - field

## CRFs & other data sources

---

- Termed 'data sources'
- Could be a faxed or scanned CRF image, PDF form data, electronic lab data, etc.
- Stored directly in the database
- Each has a unique serial number
- Each belongs to a group (could be a fax transmission, email message, upload batch)
- Each group has a serial number
- Faxes are stored in native resolution

## QC Data

---

- QCs can belong to different items: fields, CRFs, visits, subjects
- Multiple QCs per item
  - Each must have a unique type
- Multi-field QCs
  - QCs may be attached to 1+ data fields on 1+ plates
- Each QC has 2 role attributes:
  - Recipient (who will receive the QC)
  - Resolver (who is permitted to resolve the QC)
- QC reports table links QCs to the reports they are/were in
  - What has happened to the QCs in a report?

## Customizability

---

- **User-defined codes and labels for**
  - QC problem types
  - QC status codes (ultimately evaluates to one of resolved, unresolved, or notice)
  - QC resolution method
  - Missing values
  - Plate status code
  - Reason code for field or plate
  - Roles
  - Site status, region, time zone
  - etc...

## Auditing

---

- **Data changes**
  - who, what, when, why
- **Refaxed images**
  - differs from DataFax 3 notion of primary/secondary
- **QC notes**
  - added, modified, deleted
- **Study setup changes**
  - plate, module, field, visit map, centers, etc.
- **QC report history**
  - who created it, what QC notes, when transmitted
- **Task performance history**

## Double Data Entry

---

- "Heads-up" DDE
  - build sets for DDE on-the-fly
  - real time changes, 2nd reviewer wins
- DDE tracking
  - which records, when, by whom
- DDE reporting
  - summary of what has been completed
  - summary of changes made

## Refax Processing

---

- Secondary records are not stored in the database
- Old values and previous CRF images are preserved in the audit tables
- Will still be able to review prior images during duplicate resolution
- A refaxed image will be saved even if it is not made the new primary copy of the CRF

## Standard Reports

---

- Separation of content and format
- Output: XML tagged content which can be formatted using XSLT
  - deliver text, HTML, XML, PDF
- Allow new patient and visit identifiers
  - patient initials (as well as patient ID)
  - visit date (as well as visit number)
- Distribution:
  - by fax, email, URL
  - any report (not just QC reports)

## Changes since DFUG 2001

---

- Not planning to support landscape CRFs
- Subjects "belong" to a site, not a person at a site
- A module does not restrict the layout of the fields within the module. The layout may differ among different instances of the module.
- Some of the fields in a module may be hidden and assigned default values in some instances of the module.

---

## Databases

### Supported Databases

---

- Oracle
  - Implementing with 8i and 9i
- PostgreSQL
  - Implementing with 7.2
- Microsoft SQL Server
  - Implementing with 2000

## Storage Comparison

---

DataFax 3.5	56 MB
Oracle 8 (Unix)	59 MB
Oracle 9 (WinNT)	62 MB
SQL Server 2000	71 MB
PostgreSQL 7.2	74 MB

Includes 2688 primary records, images and setup  
 DataFax 4 images are scaled up to 200x200dpi

## Ease of Administration

---

PostgreSQL	<b>Easy</b> Simple command line
SQL Server	<b>Easy</b> GUI
Oracle	<b>Hard</b> Lots of tuning parameters

## Features

---

	Replication	Table space layout	Hot Backup	ODBC Interface
PostgreSQL	No	No	Yes	Yes
SQL Server	Yes	Yes	Yes	Yes
Oracle	Yes	Yes	Yes	Yes

## Costs

---

	Cost per server	Commercial Support
PostgreSQL	Free	Limited
SQL Server	\$20000	Medium
Oracle	\$40000	High

## Summary – PostgreSQL

---

- **Advantages**
  - Low (no) cost
  - Easy to administer
- **Disadvantages**
  - No advanced features like replication, table space layouts (yet)
  - Unix only

## Summary – SQL Server

---

- **Advantages**
  - Easy to administer
  - Advanced features
- **Disadvantages**
  - MS Windows only (no UNIX)
  - Cost
  - Text columns limited to 255 chars for external applications

## Summary – Oracle

---

- **Advantages**
  - Supported on both Unix and Windows
  - Many advanced features
  - Highly configurable and tunable
  - Worldwide user and training base
- **Disadvantages**
  - Cost
  - Needs Oracle DBA to administer

---

## Investigative Sites

## Centers Information

---

- Replaces DFcenters.db
- Termed 'investigative sites'
- Table for information about each site
  - Name, description, role
  - Region, time zone
  - Start date, end date, current status
  - Planned screening total, enrollment total, and enrollment rate
- Site identification aliases
- Subjects belong to investigative sites
- Includes central data management site
  - Replaces DataFax 3 error monitor
- Permissions for staff/roles at each site

## People at Centers Information

---

- Termed 'members'
- Table for information about each member
  - System name (for login purposes)
  - Formal name (for addresses)
  - Informal name (for salutations)
- Member is affiliated with one or more sites
  - Unique contact information at each site including street address, telephone number, fax number, email address
- Member fulfills one or more roles at each site
  - QC notes are directed to these roles

---

## Work Flow

---

## Work Flow

- Definition: a sequence of tasks that describe the CRF reviews.
- Process has a starting point and an end point, though there may exist many different routes.
- Different CRF pages may take different routes.

## DF 3.5 Work Flow Model

---

- Validation levels 0-7.
- Order is not enforced: user can promote and demote CRFs from any level (1-7) to any other level.
- Record status: clean, dirty, error.

## DF4 Work Flow Model

---

- Work flow tasks are defined by an administrative user.
- Tasks are linked together, in a network of task nodes.
- Order is enforced: tasks must be performed in the order specified by the work flow.
- Audit trail is kept of all tasks performed.

## “Anytime” Tasks

---

- Some tasks may be so urgent (e.g. serious adverse event) that they need to be performed immediately.
- These tasks are linked to other tasks in the work flow such that the urgent tasks may be performed at any time (order independent).

## Work Flow Routing

---

Route that CRF takes through work flow is determined by:

1. Keys: subject, plate, plate rep, visit, visit rep.
2. Investigative site.
3. User or role that performed previous task.
4. Record/CRF status: user defined status.
5. QC attributes: type, status.

## Construction of Work Flow Network

---

- Define a set of task “nodes” (squares on the work flow diagrams).
- Define a set of storage nodes (circles). A CRF resides in the storage nodes when the CRF is not in a task node.
- Connect the storage nodes to the task nodes: storage → task → storage → task → storage

## Construction of Work Flow Network

---

- Define pre-conditions that determine if a CRF is eligible for a task, e.g. John or entry\_clerks can do initial data entry; user that did initial data entry cannot do second data entry.
- Define post-conditions that determine where a CRF goes on completion of a task, e.g. after task A: plates 1 -4 go to task B, and plates 5-9 go to task C.
- Conditions are attributes of task nodes, not storage nodes.

## Building the Work Flow Network

---

- Work Flow network will be built through a graphical user interface.
- User will be able to drop task and storage nodes onto the screen.
- Clicking on a task node will bring up another dialog in which pre and post conditions are entered.

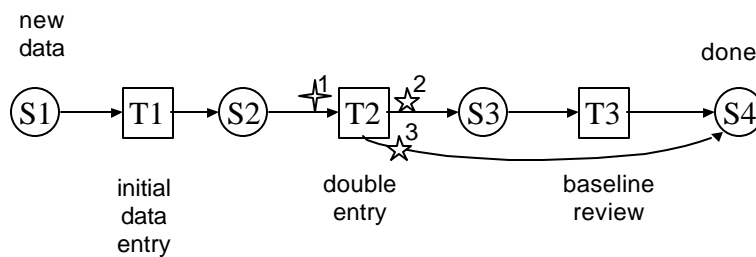
## Work Flow Example 1

---

### Tasks:

1. Initial data entry.
2. Double entry.
3. Baseline review, for visit 0 only.

## Work Flow Example 1



Preconditions:

$\uparrow^1$  must not be user who did initial data entry

Postconditions:

$\star^2$  visit 0 only

$\star^3$  not visit 0

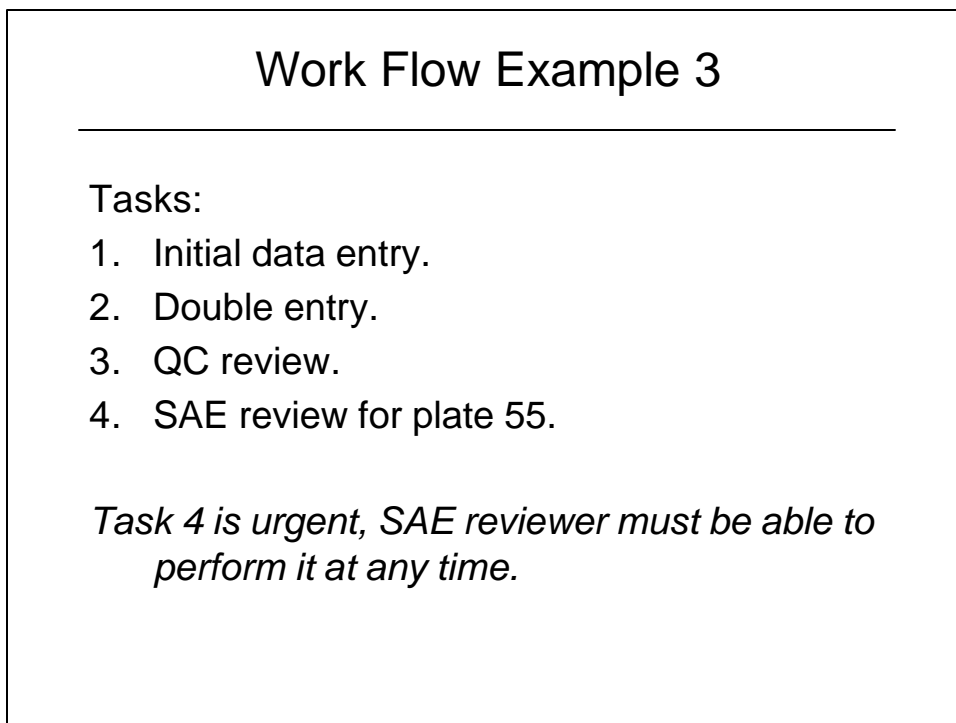
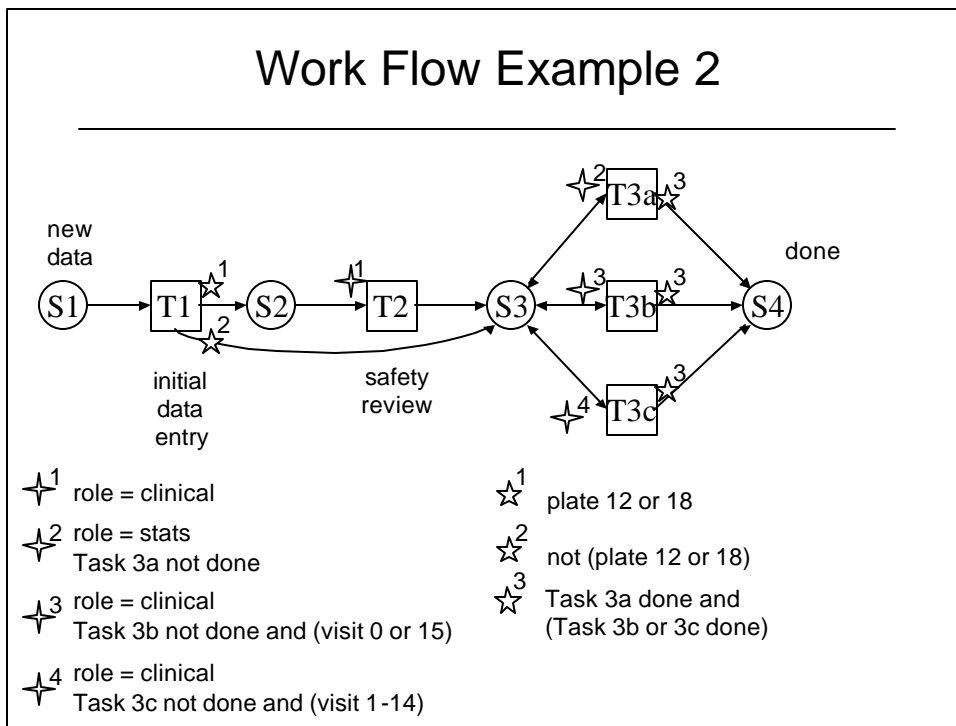
## Work Flow Example 2

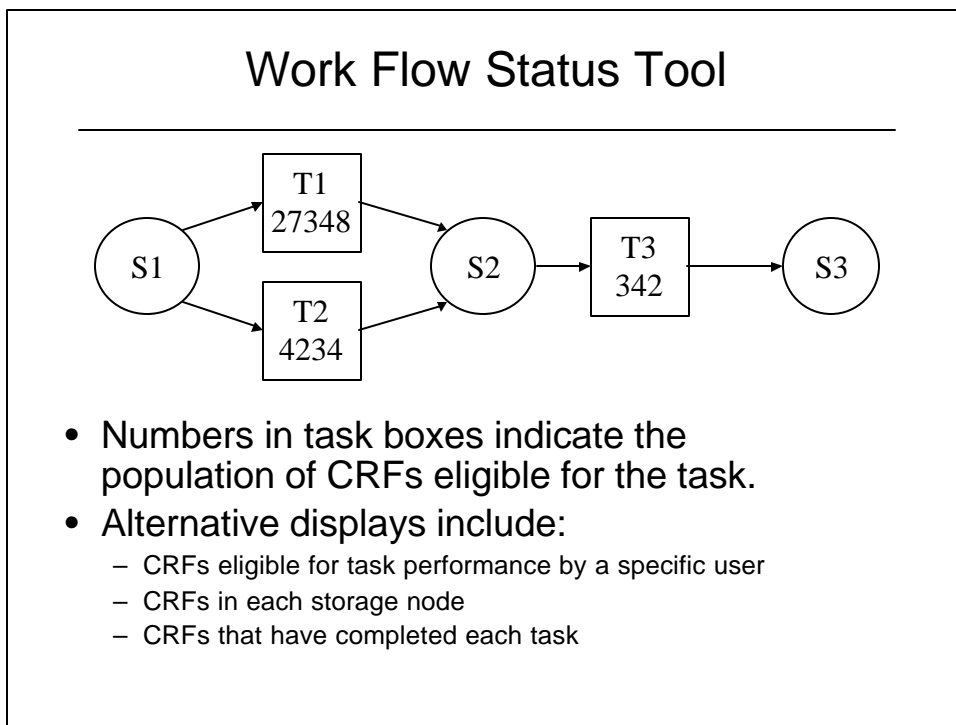
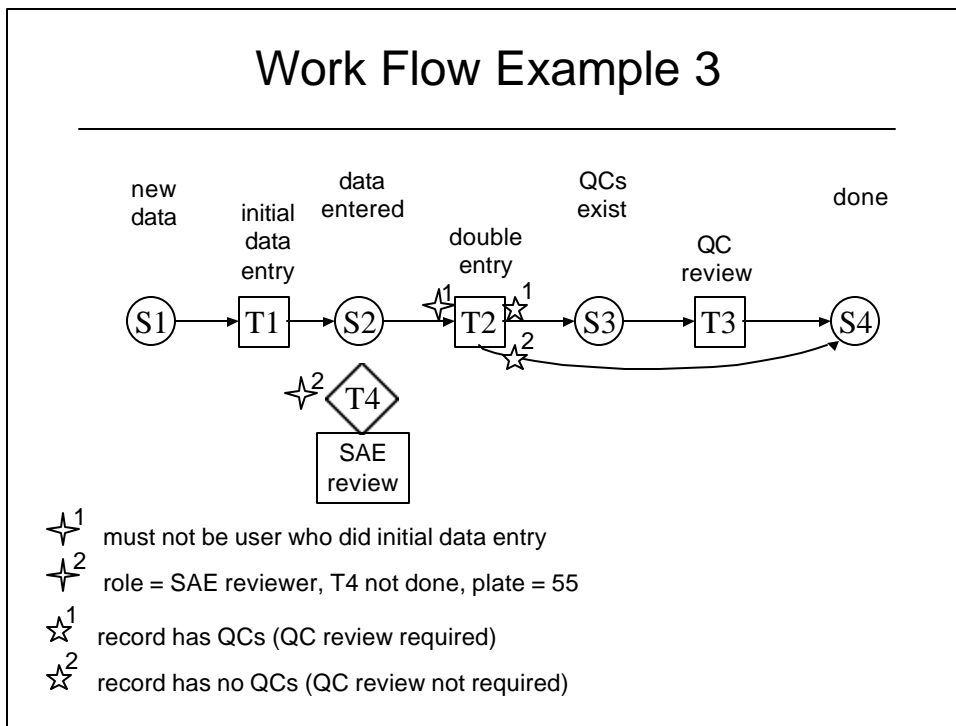
Tasks:

1. Initial data entry.
2. Safety data review for plates 12 and 18.
- 3a. Statistical review.
- 3b. Clinical review, visits 0,15.
- 3c. Clinical review, visits 1-14.

*For Task 3:*

*3a and one of 3b or 3c must be performed,  
but tasks can be done in any order.*





## What if...

---

... a change in CRF keys or other data occurs which changes the route that CRF should take through the work flow?

- After each task is performed, a check is made to ensure that the CRF still has a route to the task that was just performed.
- If not, the CRF is put into the first storage node in the work flow for which it has an eligible down stream task.

## What if...

---

... a CRF reaches a dead end in the work flow, either because, for the task just completed:

- it does not meet any of the post-conditions
- it meets more than 1 of the post-conditions

and consequently the next storage location can not be determined?

- CRF will be put into a special storage node for CRFs that are unable to advance – “limbo”.
- Administrator will have to edit work flow conditions to let the CRF advance.

## What if...

---

... a new task must be added to the work flow after some CRFs have been processed?

- Addition of a new branch is easy (e.g. a special review of new CRFs for a side study).
- Insertion of a new task between two others:
  - move eligible CRFs to storage node preceding new task
  - if desired, add post-conditions so these CRFs can skip over completed tasks that are down stream of the new task.

## DF4 Work Flow Model

---

### PROS:

- Work flow is enforced by the software.
- Tracing CRFs through workflow is improved.
- Easier and more accurate performance reports showing number of CRFs reviewed at each task by each reviewer.

### CONS:

- Work flow bottlenecks may arise.
- Work flow must be designed and set-up.

---

## CRF Review Tool

---

### General Features

- All-in-one tool
  - Single access point for all CRF review functionality
  - incorporates DFstatus, DFvalidate, DFviewer, DFidl, DFqc
- Multiple instances (MDI)
- Movable and dockable windows
- Mouse and keyboard equivalents for every action
- Minimize number of dialogs

## Features

---

- **Switchable views**
  - Top and bottom (ala DFvalidate)
  - Left and right (ala DFlite and DFviewer)
  - Overlay
- **Tabular data entry**
- **Zoomable main window**
- **Fields are adorned with icons to identify attributes**
  - QC notes, missing data, comment, reason, history
- **Integrated with workflow**

## Overlay view

---

- **Data entry for current field appears at its position on the CRF image or other data source**
- **Display of current data for other fields can be toggled**
- **Supplemented by tabular data entry**
- **Permits viewing of greater portion of data source**
- **Reduces eye movement**

## CRF selection

---

- For new CRFs
  - By data source
  - By site alias
- For previously reviewed CRFs
  - By key fields
  - By query – includes query builder
  - By example
  - By task
  - Request supplementary CRFs in addition to CRFs requested in selection

---

## Progress Report

## Current Status

---

- Not as far along as we'd like to be
- Big project
- Database design is critical
- Building a validated system multiplies the effort needed
- Building a new system to exceed the expectations of an existing, mature system is hard

## Current Status

---

- Database design and implementation is 95% complete
- CRF review tool is 40% complete
- Setup tool is 25% complete
- Work flow is 50% complete
- New data types: 90% complete
  
- Estimate 5 person years work remaining

## CDSI Objectives

---

- Beta available in mid 2003
  - Including core functionality for database definition, all CRF and data handling, and workflow
- Add remaining important features in 2<sup>nd</sup> half of 2003
- Product introduction and initiation sessions following DFUG in 2004
- Continue support for DataFax 3 until no longer required

## Issues for Discussion

---

- Migration from DataFax 3
- Command-line interfaces
- Test and production databases
- Conflict between CRF-level action and module-level permissions