Centralizing the Mink Survey at the National Agricultural Statistics Service

Roger Schou & Emily Caron
National Agricultural Statistics Service
IBUC XIII
Baltimore, Maryland USA
The Ghost of CASIC Past - Dissemination

- **Menu System**
  - VB 6
  - Numerous .exe files (12 frequently used)
  - LAN updates required in each Field Office

- **Decentralized Blaise**
  - Blaise 4.80
  - Up to 43 Blaise datasets for a survey
  - No national-level survey administration tools
The Ghost of CASIC Past – Dissemination (cont.)

• Six Data Collection Centers
  – Calling is more centralized, but Blaise datasets are still disseminated
  – Continuous physical movement of Blaise datasets to and from Data Collection Centers
  – Still no national reports
The Ghost of CASIC Present - Centralization

- April 2010 – Mink Survey Conducted
  - All data collected and edited in Blaise in two states (Utah & Wisconsin)
  - Small sample size (415)
  - Longer data collection window
  - Used Blaise 4.82 build 1529
MySQL Database

- Central database in MySQL
- Generic .BOI files used
- In-depth data partition type
- Versioning was turned on
CASIC Menu

- Visual Basic .NET
  - One set of source code
  - Can be updated without LAN Updates
- WEB-enabled
- Dynamic - based on user’s role and state’s responsibilities
- Reads the Windows AD table
Client Requirements

- .NET Framework had to be installed
- MySQL Driver had to be installed
- ODBC System data sources had to be set up to provide connection info to MySQL
- Eight Blaise .DLL’s had to be registered
- All of this was pushed out prior to the survey
CASIC Tables

- MySQL tables in addition to the 8 generic tables created by Blaise
  - CASIC_SurveyInfo
  - CASIC_FAT
  - CASIC_Management
CASIC_SurveyInfo

- Critical info about each survey
- Survey code, year, month, and day
- Instrument name, data collection dates
- Survey type and other indicators to aid in managing the survey
- Allows selective display of surveys based on user and date
- Info used by the survey menu
CASIC_FAT

- FIPS Assignment Table
- Controls who sees which surveys as well as certain areas of the menus, based on their role
  - Responsible for data collection
  - Responsible for editing
  - State which “owns” the data
CASIC_Management

- Fields within the instrument that will be used to filter or sort the dataset in Manipula or the Data Center
  - DCCFips, ECFips, FOFips, Batch, ProcessSwitch, State, and County
- All fields are indexed
Issues Encountered

CATI Service Errors

• Large number of CATI Service Errors
  – Initially solved by downing the service
  – After communicating with our database experts and members of the Blaise Team at Stat Neth, we changed some ODBC settings
  – Also got a new version of Blaise
  – This drastically reduced the number of CATI Service errors
Issues Encountered
Performance

• Performance did not start out as ideal
  – Originally, Blaise software was accessed from the central application server
  – Part way through, we began accessing it from the local FO LAN.
  – Changing the location and the updated version both helped with the performance.
Issues Encountered
Time Zones

• The central database is physically located in the Eastern Time Zone
  – Therefore, the “home” time zone for all states is Eastern Time Zone
  – All states will use the same time zone differentials
Conversion Schedule

• Beginning in November 2010 with one survey
• Continue adding surveys through December 2012
The Ghost of CASIC Yet to Come

• Citrix
  – All Virtual Workstations by summer of 2011
  – Eliminates the need for specialized workstation installs (DLLs, NET framework)

• No Versioning to keep the database smaller
  – ETL to store all iterations of the data

• Insure proper system failover is in place
The Ghost of CASIC Yet to Come (cont.)

- ETL process to pull data from MySQL database to Work In Progress (WIP) database
- UDL file to accommodate the requirement of the password changing every 30 days
- The use of CRON jobs to do some of the processing overnight
- One .BTR file versus multiple .BTR files
Conclusion

- Centralization is moving forward
- National-Level Reports
- Other systems are centralizing
- Operational Efficiency is key
Questions