



**N A S S**

---

FACT FINDERS FOR AGRICULTURE  
UNITED STATES DEPARTMENT OF AGRICULTURE

# **Implementation of a New Processing System for the 2002 Census of Agriculture: Mixed Results**

Oct. 20, 2003

# The Good News

- ✦ We implemented a completely new processing system for the census, which included extensive micro- and macro-editing capabilities.
- ✦ We used scanning and OCR/ICR data capture methodology for the first time – and on a very grand scale!
- ✦ We implemented our first major survey processing system utilizing database technology.
- ✦ As an Agency we'll survive what was a very ambitious effort – changing everything in one cycle.

# The Bad News

- ✦ The developmental and implementation process was very rocky -- plagued by technical and administrative issues.
- ✦ The final processing system ended up much different (and compromised) from the one initially envisioned.
- ✦ Lack of testing time and test data resulted in numerous system bugs being discovered in production.
- ✦ The targeted rate for the manual review of records was 10 percent. The actual rate was several times that.
- ✦ Timelines slipped badly, and we are still struggling to clean up the data.

# Technical Issues

- ✦ The planned implementation of AGGIES-based error localization was scrapped for processing efficiency.
- ✦ The resulting DLT-based editing modules (with donor imputation) were still too slow.
- ✦ Data bases for administrative information (Sybase) and actual data (Red Brick) were frequently out-of-sync.
- ✦ Synchronization issues resulted in erroneous data entering donor pools. These were then propagated to other records.
- ✦ Scanning for image was an unqualified success, but OCR/ICR data capture was not. Clean-up of incorrectly captured data significantly taxed Agency staff.

# Administrative/Structural Issues

- ✚ Merging the census responsibilities into the Agency's functional structure resulted in significant delays in preparing for the 2002 census.
- ✚ Staff's collateral duties with the Agency's on-going survey program (shorter-term needs) always took priority over census preparation. Fire-fighting! Census wasn't considered a fire until it engulfed us.
- ✚ Matrix-management issues. Cross-divisional teams in a line management structure. Whose priorities come first – your supervisor's or your team leader's?

# Key Lessons Learned

## (Courtesy of the School of Hard Knocks)

- ✦ Start preparing for the next census much earlier!!
- ✦ Consider alternative structural approaches that will give more priority to census preparations, by reducing collateral duty conflicts.
- ✦ Settle on one database for both administrative information and data.
- ✦ Implement procedures to avoid the OCR/ICR capture of erroneous data in crossed-out sections of the questionnaire.

# Current Status/Plans

- ✚ While continuing to clean up the 2002 census data, we're already looking ahead toward 2007.
- ✚ We'll research alternative editing/imputation approaches, including Statistics Canada's NIM, over the next two years.
- ✚ Recommendations for the 2007 census will be made by the end of 2005 – in time for system development, test data generation and testing.