

Alex Roberts, Assistant Director of Data Processing, NYS Division of Criminal Justice Services

Alex Roberts talked about the Division of Criminal Justice Services (DCJS) experience in redesigning and converting their Computerized Criminal History (CCH) system. This large database contains 9.1 million histories that are made up of 13 million criminal cycles. It has a growth of an average of 3000 histories/day. Therefore, the project is mainly a computer systems migration project, involving the redesign of a very large mainframe-based, on line system to one that uses object-oriented design, 3-tier client/server architecture and relational database design.

The major principles that they applied to the data design were the following:

1. Storing all data in standard format unless there is a technical impediment to doing so. In that case, the data would be stored in a way that can ensure data standards transmission compliance.
2. Translating DCJS codes into standard codes and eliminate the DCJS codes. When the data standard elements are inadequate, they would be stored in a format that is conducive to good data modeling principles. Where there are conflicting standards (criminal justice versus statewide), the standard that is most beneficial would be used. An ability to transmit in either standard is required.
3. Documenting all changes to the way current data are stored. If the decision were made not to store in data standards, the reasons would also be documented.

The problems that they encountered dealt with the conversion of data to comply with established data standards; the conversion of invalid data, and the conversion of incorrect data (data that passes constraint edits, but is not correct). The successes were the following: (1) development and adherence to a data standards policy; (2) only one element (DOB) with invalid dates was converted to new database using 2 data elements; (3) resources were allocated to clean up the most serious problems in the conversion of incorrect data, mainly by researching source documents.

The timeline of the project:

- 2 years of database design, and redesign,
- 1 year of analysis and coding of the conversion program,
- multiple cycles of test conversions,
- projected 4 months conversion run,
- the 2 databases must continue to run in parallel for several years while the migration project is completed.

Alex ended his presentation by giving a few recommendations:

- Develop a thorough understanding of the data. Review the current processing. Talk to people who use the data;
- Don't be seduced by development speed. Take time and care in the analysis, design, testing, and tuning;
- Never misuse data; each data element should have only one definition.