Information Quality at the Transportation Security Administration

July, 2010

Introduction

• Why information quality is so important from the perspective of the Performance Management Information System (PMIS) program.

• What is being done to ensure info quality.

• What is working well.
Introduction

“Scientific management means a constant search for the facts, the true actualities, and their intelligent, unprejudiced analysis.”

- Alfred P. Sloan
  MIT Class of 1895

Overview of PMIS

- **Performance Measurement Information System (PMIS):** The Performance Measurement Information System is a web-based application used to collect TSA metrics and measures. Data entry is accomplished by web browser or hand-held device. PMIS directly supports TSA’s goal of becoming a performance-based organization.

- **Business Intelligence (BI) Tool:** The Performance Information Management System (PIMS) is a state-of-the-art tool which supports analysis, dashboarding, graphing, and reporting from a consolidated data warehouse. PIMS today reports on data from 18 TSA data sources.

- **Airport Information Management (AIM):** Next-generation data entry system to assist airports in managing their day-to-day business. Provides airports a consistent national interface with headquarters’ support and ports existing field applications to this national enterprise application. Supports real-time business intelligence reporting.
Security and data quality layers

Terrorist Paths
- Intelligence
- Counterterrorism Task Force
- Mitigate Unlawful Helper
- Cow Tracking
- VIP
- No Fly List and Passenger Pre-Screening
- Behavior Detection Officers
- Crew Vetting
- VIPR Canines
- Transportation Security Inspectors
- Checked Baggage Random Employee Screening
- Bomb Appraisal Officers
- Federal Air Marshal Service
- Federal Flight Deck Officers
- Trained Flight Crew
- Law Enforcement Officers
- Hardened Cockpit Door
- Passengers

Raw Data
- Cleansing
- Data Validation
- Data Integration
- Standardization
- Change Control
- Flagging and Review
- Data Boundaries
- Real-time Reporting
- Analytical Reporting
- Consumer Reports

Accurate Intelligence

Integrated Data Sources - PMIS
- PMIS
- OCR
- BTS
- SPOT**
- Management Objective Reports
- Integrated dashboards
- Email Alerts
- Mobile Devices

**Now in PMIS
Custom hierarchies with integrated data collection and reporting.

Data collection methodology:
- Mapped to business process
- Requirements include data values consistent within multiple systems
- Working with master data

One “safe source” of operational data shared across multiple applications and program offices
Testing (Quality Control)

<table>
<thead>
<tr>
<th>Work Product Category</th>
<th>Tool(s)</th>
<th>Verification Method(s)</th>
<th>Validation Method(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Documentation</td>
<td>Rational ClearQuest</td>
<td>Peer Reviews</td>
<td>Formal Reviews, QA Audits, Customer Acceptance of documents.</td>
</tr>
</tbody>
</table>

Challenge – inconsistent geographic data

Many TSA systems allow user-entered checkpoint names, machine information, and airport location data. These text-based entries may differ depending on the individual(s) performing data input.

Furthermore, changes to geographic information must be propagated in a timely manner to both data entry and reporting systems.
Consistent hierarchies for data collection and reporting across PMIS/PIMS/AIM. All geographic values tagged with unique identifier and synchronized via ETL.

The BI Reporting system is also designed to accommodate languages and airports from across the world.
Challenge – Data quality and privacy across systems

TSA systems do not employ a standard unique identifier for each user. As such, each application may utilize a different solution for user authentication.

Examples of system identifiers:
• Arbitrary auto-generated integer value
• Social Security Number
• email address
• Network login ID
• Last four digits of social + first two letters of last name
• First letter of first name + last name

Solution – cross-system identifier reference

Master reference table containing system identifiers from multiple applications. PII is encrypted using industry-standard algorithm

<table>
<thead>
<tr>
<th>SSN</th>
<th>Name</th>
<th>Hours</th>
<th>Dollars</th>
</tr>
</thead>
<tbody>
<tr>
<td>123456789</td>
<td>John Doe</td>
<td>40</td>
<td>$1000</td>
</tr>
<tr>
<td>987654321</td>
<td>Jane Smith</td>
<td>80</td>
<td>$2000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SSN</th>
<th>Name</th>
<th>Courses taken</th>
<th>Average Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>345345078</td>
<td>Bob Smith</td>
<td>6</td>
<td>90.5</td>
</tr>
<tr>
<td>873462961</td>
<td>Robert Doe</td>
<td>15</td>
<td>88.1</td>
</tr>
</tbody>
</table>

Cryptographic Function
Challenge – Data entry mistakes

Users may accidentally enter data values that are higher or lower than intended. These values should be flagged for review prior to being reported nation-wide.

Solution – Boundary definition

High and low boundary limits can be set on a per-metric, per-category basis for PMIS submissions. Submissions with values outside of boundaries are automatically flagged.
Solution - Flagged submission review

Each submission can be individually reviewed for accuracy prior to being “promoted” for reporting in the Business Intelligence tool.

Challenge – inconsistent item category data

Many TSA systems allow user-entered item names, item groups, and item category data. These text-based entries may differ depending on the individual(s) performing data input.

Furthermore, changes to item information must be propagated in a timely manner to both data entry and reporting systems.
Data Entry – Consistency and Control

Drop-down lists and data entry screens can be customized according to individual airports’ needs while maintaining standard category definitions.

Challenge – data feed monitoring

The TSA business intelligence tool receives daily data feeds from a number of source systems. Each system may experience outages, maintenance, or service interruption during its scheduled data load.

Users at headquarters and in the field must be aware of any data load failures prior to running operational reports the following morning.
Daily email distributions indicating success or failure of various data interconnections.

Submission summary allows airports to “drill down” for details about each day’s data.
Application utilizes asynchronous JavaScript and XML (AJAX) to conduct data manipulation without repeated database calls. This reduces server traffic and database load.

The TSA business intelligence tool receives daily data feeds from a number of source systems. Data from these systems may be up to 24 hours old in the data warehouse.

Users at headquarters and in the field have a need for more timely reporting and alerts.
Real-time data entry and reporting integration exists between AIM and PIMS utilizing a shared warehouse for “one safe source” of operational data. Referential data is synchronized in real time.

Event- or time-based alert capability from integrated data warehouse.
Integration options

Each program office can be granted access to different areas of the TSA BI data warehouse. This flexibility enables each business unit the ability to control its own data while pushing select elements agency-wide.

Option 1 – Standard Development Process
BI analytical team creates and maintains all publicly-available reports for a program office; users are allowed to create private “ad-hoc” reports within the application.
Integration options

Example developer-created “Amount of Money Left at Checkpoint” analysis report showing values for a specific date range and subset of airports.

Integration options

Option 2 (currently in use) – Collaborative Development Process
BI analytical team and program office developers create and maintain all publicly-available reports for a program office; users are allowed to create private “ad-hoc” reports within the application.
Example user-created “what-if” analysis report showing DCA throughput vs. capacity values for specific dates.

**Option 3 – Program Office Development Process**
Program office team creates and maintains all publicly-available reports for a program office; users are allowed to create private “ad-hoc” reports within the application.
Integration options

Example user-created report showing machine maintenance status and outages.

Media Comments

“The focus on metrics allows the TSA to avoid getting ‘hamstrung’ because it doesn’t rely on intuition to make decisions or have information stored in disparate spreadsheets, data marts and reports.”

- Wayne Eckerson
  Director of research and services
  The Data Warehousing Institute

“PIMS tracks more than 1,000 metrics in total, but TSA has customized dashboards and reports for approximately 30 roles within the agency, including top executives, optimization teams (they’re the BI power users), and front-line security personnel.”

- Special Report: Business Intelligence Gets Smart
  Intelligence Enterprise
  September 2008

“If your most recent trek through airport security was less painful than even a year or so ago, thank the Transportation Security Administration’s overarching business intelligence project, dubbed the Performance Information Management System, or PIMS.”

- BI Efforts Take Flight
  InformationWeek
  10/13/2008