The Future of Tax Data Warehousing: Moving Towards a Multi-Purpose, Revenue Agency-Wide Data Warehouse

Presented by:
Paul Morris
Revenue Solutions, Inc.

July 12, 2005

Agenda

- **Current Context** for Data Warehouse Projects – Primarily “Single-Purpose”
- Vision of **Broader Functionality**:
  - Multi-Purpose
  - Agency-Wide
- **Framework** for Data Warehouse Planning, Implementation and Evolution
- Massachusetts DOR **Case Study**
Central Themes

- The **Foundation** is More Important than Specific End Uses and Tools
- The Core Design Must be Sound, and then **Grow with the End in Mind**
- The **Functionality** Supported by Data Warehouses will **Continue to Expand**
- The Data Warehouse will Enable Future Business Improvements in **Productivity, Compliance, Revenue Analysis, Performance Measurement and more.**

Part 1

The **Current Context** for Data Warehousing
Current Context

Shift in an Agency’s Primary Challenge from *Data Processing* to *Data Analytics*
- Increasing Self-Service Channels (Phone, Web)
- Increasing Process Automation
- Data Explosion
- OLTP Capabilities have Improved…

So what’s next?

Use the Data to…
- Drive Improvements in
  - Productivity,
  - Revenue Management,
  - Operational Efficiency, and
  - Taxpayer Compliance
- With Available - Often Fewer - Resources

Data Trends

<table>
<thead>
<tr>
<th>QUALITY</th>
<th>Captured and Validated Automatically on the Front End Before Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>QUANTITY</td>
<td>Increased Number of Interactions and “Events” Recorded for Each Taxpayer</td>
</tr>
<tr>
<td>COMPLETENESS</td>
<td>Approaching 100% Electronic Data Capture (No Keying) Coupled with Relatively Inexpensive Data Storage</td>
</tr>
<tr>
<td>SOURCES</td>
<td>Expanding Data Sharing Programs Between Agencies, States and IRS</td>
</tr>
</tbody>
</table>

Result: Data is more valuable overall … *if it is used effectively*
What’s Next?

Data Warehouses Support:
- Non-Compliance Discovery
- Data Analytics and Predictive Modeling
- Applying Consolidated Taxpayer Portfolios when Making Decisions
- Integrated Compliance Program Management
- "Lifecycle" View of Effectiveness Measures
- Revenue Forecasting and Analysis
- Operational Reporting and Tracking Against Performance Benchmarks
- More...

Expectations are High…
The Sky is the Limit?

Part 2

**Broader Functionality:**
Moving Towards a Multi-Purpose and Agency-Wide Data Warehouse
Non-Compliance Discovery

- Matching Internal Filing Records with External Data Sources (e.g., IRS, Customs) to Identify Non-Compliance:
  - Non-Registrants
  - Non-Filers
  - Deficiencies (Desk Audit)
  - Calculation of Tax Assessments
- Apply Data Mining Techniques
- Issue Notices and Bills

Build Taxpayer Portfolios... Then “Mine” for Non-Compliance

Predictive Modeling

- Predictive Modeling:
  - Use Data from Past Audits to Construct Models
  - Validate on “Holdout” Sample of Past Audits
  - Apply Models to Select and Prioritize New Leads
  - Confirm Effectiveness and Adjust Model Over Time

Decision Point:
- Should an Audit be Conducted?
- What is the Scope of the Audit?
- What form of Audit?

Predictive Modeling Can Also Be Used For:
- Collections Treatment
- Taxpayer Education Initiatives
- Non-Compliance Discovery
Compliance Program Management

- Forecast Revenues for Projects
- Accurately Plan Work Based on Prior History
  - Inbound Calls
  - Collections
  - Fly Backs (RPO)
  - Abatement Requests
- More Immediate Feedback of Program Results
- Shift Resources to the *Effective Programs and Non-Compliant Segments*

Taxpayer Portfolios – Decision Support

- Taxpayer Portfolios Provide a “Single View” for Making Decisions
  - When and How to Provide Education
  - How to Collect
  - When and How to Audit
- Improved Customer Service
- Integrated Compliance and Education Activities – One Contact, One “Message”
Revenue Analysis

• Provide Analysts Access to Raw and Aggregated Data
• Analyze Revenue Trends and Explore Data with Drilldown/OLAP Functions
• Conduct “What If” Analysis to Determine Impact of Tax Changes on Revenues

Reporting & Ad Hoc Query

• “Unlocking” the Data
• Making Queries Accessible to End Users
• Web-deployed Presentation Tools
• Additional Aggregate and Summary Data Structures are Required
Scope of the Warehouse

*From the Discipline of Systems Science:*

The optimum solution to a problem is often larger than the sum of the component parts.

It requires a widening of the problem space to appreciate a common, shared solution.

Part 3

A *Framework* for Data Warehouse Implementations
## Warehouse Framework

- **Data Sources**
- **Assemble the Data**
- **Aggregate/Sub-Divide the Data**
- **Data Warehouse**
- **Data Marts — Queried Data**
- **Tools and Interfaces**
  - "End Users"
  - External and Internal
  - Tape-Based or Online
  - Refresh Strategies
  - Regular Interfaces
  - Validation
  - Standardization
  - Matching Engine
  - Build/Manage Portfolios
  - Common Data Model
  - Normalized
  - Tuned for Inserts and Updates
  - Aggregation Engines
  - Derived/Complex Attributes for Modeling
  - Populate/Refresh
  - Pre-Aggregated Data
  - De-Normalized Views
  - Specialized Marts
  - SQL-Compliant Tools
  - Web-based Delivery
  - OLAP/Report Engines
  - Analytical Tools
  - De-Normalized Views
  - Specialized Marts
  - Attributes for Modeling
  - Populate/Refresh
  - Tuned for Inserts and Updates
  - Matching Engine
  - Build/Manage Portfolios
  - Regular Interfaces

## Tax Data Warehouse Framework

- **OLAP**
- **OLAP Cubes**
- **Ad Hoc Query and Reporting**
- **Reporting Data Marts**
- **Analytical Tools**
- **Tax Changes Analysis Data Mart**
- **Data Mining**
- **Predictive Modeling**
- **Audit, Collections, Discovery Analysis Data Mart**
- **Aggregation/Refresh Engines**
- **Data Warehouse**
- **Portfolio Viewer**
- **Calculate Tax Assessments**
- **Calculate Model Scores**
- **Portfolio Matching & Building Engines**
- **Extract, Transform, Load**
- **Query Warehouse**
- **Ad Hoc Query**
- **Compliance Programs and Decision Engine**
- **Web Systems**
- **External Tapes**
- **Legacy Systems**
Key Principles

- Get the Foundation Right – All Functionality Requires Accuracy, Currency and Availability
- Getting Value from Business Intelligence Tools Often Requires “Back-end” Work
  - De-normalized Data Structures
  - OLAP Structures
  - Performance Tuning
- Design and Plan for Expansion
  - Extensible Data Model
  - Add New Sources Over Time

Part 4

A Case Study:
Massachusetts Department of Revenue
**MA DOR Non-Compliance Programs**

**Business Process Flow:**

- Calculate Tax (Non-Filer or Deficiency)
- Send Initial Letter (Removed for Most)
- Send Bill(s)
- Support Inbound Calls and Letters
- Load Data and Build Portfolios
- Follow-up Bills with Collection Activities
- Define “Sub-Populations”
- Monitor Project Effectiveness – Apply Feedback

**Table:**

<table>
<thead>
<tr>
<th>Project Year</th>
<th>Notices Issued</th>
</tr>
</thead>
<tbody>
<tr>
<td>TY2000</td>
<td>31,000</td>
</tr>
<tr>
<td>TY2001</td>
<td>47,000</td>
</tr>
</tbody>
</table>

**Warehouse-Driven Non-Compliance Program Totals**

<table>
<thead>
<tr>
<th>Year</th>
<th>Revenues</th>
</tr>
</thead>
<tbody>
<tr>
<td>FY2003</td>
<td>$9 million</td>
</tr>
<tr>
<td>FY2004</td>
<td>$27 million</td>
</tr>
<tr>
<td>FY2005</td>
<td>$58 million</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td><strong>$94 million</strong></td>
</tr>
</tbody>
</table>

**Refund Rate:**

- 3.2%
- 38.6%

**Net Collections:**

- After the first 5 months of each Individual Income Non-Filer project... $1,282,000
- TY2001: $1,325,000

**Gross Collections Notices Issued:**

- TY2000: $841,000
- TY2001: $58 million

**FY2003 Revenues:**

- $9 million

**FY2004 Revenues:**

- $27 million

**FY2005 Revenues:**

- $58 million

**FY2005 Revenues:**

- $9 million

**MA DOR EDW Implementation**

- Massachusetts DOR Economic Data Warehouse (EDW) Implementation:
  - Conversion of 5+ Years of Transaction History from the MASSTAX Legacy System
  - Daily Interface of Legacy Financials, Revenue Accounting, and Taxpayer Registration Updates
  - Advanced Revenue Analysis
    - OLAP Cubes Provide Functionality to “Drill” into Data to Explore and Isolate Changes
    - Apply Tax Calculator and Create a Results Cube for “What If?” Analysis of Tax Changes
Data Warehouse Lessons Learned

- Success Requires Executive Commitment
- Build the Warehouse and Deliver Functionality in Stages – Within a Flexible Design
- Focus on the Foundation
  - Data Model that Accommodates Agency-Specific Tax Data and Various External Sources
  - Data Quality on the Front-End – Before Loading
  - Matching and Portfolio Building
- Implement Quality Assurance Processes
  - Testing Selection Routines, Tax Calculations, Interfaces
  - Testing “Synchronization” Between Tax System and Data Warehouse

Revenue Solutions, Inc.

**Mission Statement**

“Assist revenue agencies to maximize collections, increase compliance, improve customer service and streamline operations through the use of enabling technologies, in particular, integrated tax and tax data warehouse solutions”

- Incorporated in May, 1996
- Headquartered in Pembroke, MA with a DiscoverTax® Solution Center in Sacramento, CA
- Dedicated exclusively to providing products & services to tax agencies
- Approximately 90 consultants with over 450 combined years of revenue systems consulting
- Deep tax administration domain expertise

www.RevenueSolutionsInc.com
(888) 826-1541
Questions & Answers