Using Data Quality Methods at the Federal Railroad Administration

Improving the Archiving and Retrieval of Safety Information

Agenda

- Introduction
- Background
- The Context of the FRA Enterprise
- Data Quality
- Information As Product
- Information Product Maps
- Conclusions

5/27/2008   **2**   Federal Railroad Administration

Introduction

- **Mission**
  - □ *Ensure the safety and efficiency of passenger and rail freight services in the United States.*
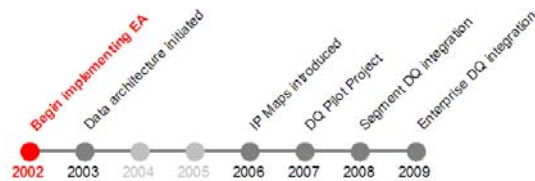
- **Task**
  - □ *500 Inspectors*
  - □ *Inspect 230,000 miles track, 1.2 million freight cars, 20,000 locomotives, and 89,000 track miles of signal and train controls.*

- **Challenge**
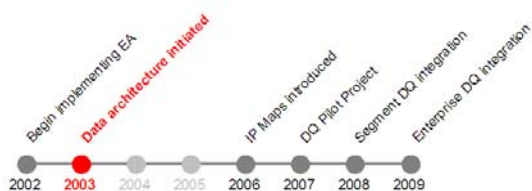 - □ *Optimize efficiency while guaranteeing quality*

5/27/2008   **3**   Federal Railroad Administration

Timeline **2002**

- FRA began implementing EA In 2002.

  - □ *Eight functional areas*
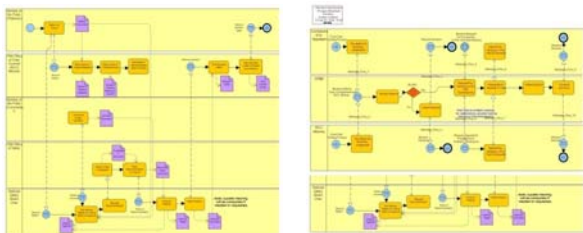 - □ *Three view levels*
  - □ *First cut, since refined.*

5/27/2008   **4**   Federal Railroad Administration
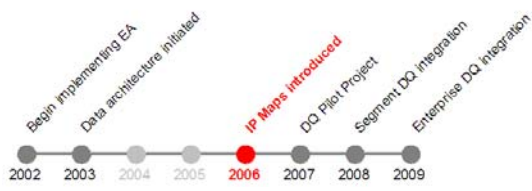
Timeline **2003**

- Data architecture initiated in 2003
  - *Captures data exchanges within and between LOBs using both traditional and object-oriented techniques.*
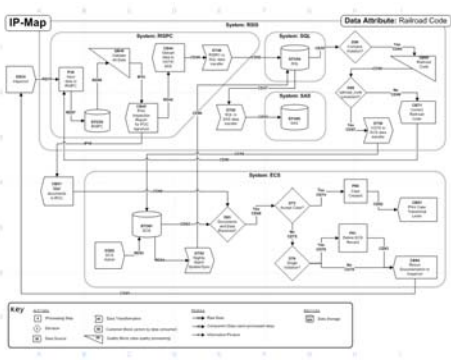
5/27/2008   **5**   Federal Railroad Administration

Timeline **2006**

Wang Information Product Map (IP Map) methodology introduced in 2006.

5/27/2008   **6**   Federal Railroad Administration

Timeline **2007**

Begin implementing EA
Data architecture initiated
IP Maps Introduced
DQ Pilot Project
Segment DQ Integration
Enterprise DQ Integration

2002 2003 2004 2005 2006 **2007** 2008 2009

- Data Quality Pilot Project conducted in Fall 2007.

- Focus of project on eight essential data attributes shared between two LOBs.
  - *Office of Chief Counsel (RCC)*
  - *Office of Safety (RRS).*

5/27/2008   **7**   Federal Railroad Administration

Timeline **2008**

Begin implementing EA
Data architecture initiated
IP Maps Introduced
DQ Pilot Project
Segment DQ Integration
Enterprise DQ Integration

2002 2003 2004 2005 2006 2007 **2008** 2009

- Segment-by-segment analysis of FRA enterprise throughout 2008.
  - *New Initiatives*
    - Disaggregation and cataloging of institutional reports and data calls.
    - Coordination, cooperation, and data sharing between EA and Security data collection teams.

5/27/2008   **8**   Federal Railroad Administration

Timeline **2009**

Begin implementing EA
Data architecture initiated
IP Maps introduced
DQ Pilot Project
Segment DQ integration
Enterprise DQ integration

2002   2003   2004   2005   2006   2007   2008   2009

Database consolidation and normalization

□ *Rationalized data structure as keystone of Data Quality Program and Service Oriented Architecture (SOA).*

5/27/2008   **9**   Federal Railroad Administration

The Context of the Enterprise

■ **Five Lines of Business (LOB)**

□ *Office of the Administrator*

□ *Office of Safety*

□ *Office of Railroad Development*

□ *Office of the Chief Counsel*

□ *Office of Financial Management and Administration*

5/27/2008   **10**   Federal Railroad Administration

Data Quality

"Fitness for Use"

Data Quality (cont.)

Information Quality Category	Information Quality Dimensions
Intrinsic Information Quality	Accuracy, Objectivity, Believability, Reputation
Accessibility Information Quality	Accessibility, Access Security
Contextual Information Quality	Relevancy, Value-Added, Timeliness, Completeness, Amount of Information
Representational Information Quality	Interpretability, ease of understanding, concise representation, ease of manipulation

Information As Product

- Product vs. By-Product
  - *Consumer vs. Producer measures of quality.*

- Formal recognition of the criticality of data quality to mission.
  - *Organizational structures, positions, and processes to manage data quality.*

5/27/2008   **13**   Federal Railroad Administration

Information As Product (cont.)

	Information as Product	Information as By-product
What is managed?	Information; information product life cycle	Hardware and software; systems life cycle
How is it managed?	Integrated, cross-functional approach that encompasses information collectors, custodians, and consumers	Integrate stove-pipe systems; control individual components; control costs
Why manage it?	Deliver high-quality information to consumers	Implement high-quality hardware and software system
What is success?	Deliver high-quality information continuously over the product life cycle; no GIGO (garbage in, garbage out)	The system works; no bugs
Who manages it?	Chief information officer (CIO); information product manager	CIO; information technology director and database administrators

5/27/2008   **14**   Federal Railroad Administration
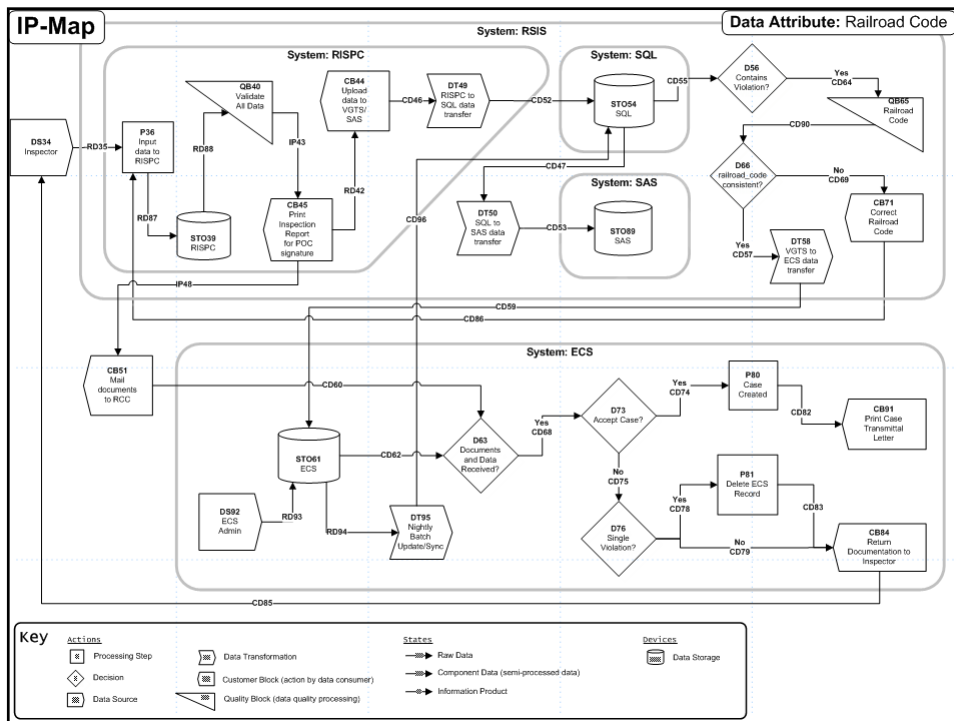
Information Product Maps (IP Maps)

- Designed to graphically describe data workflows and transformations.
  - *Can represent single data elements or logical groupings of multiple data elements.*

- Identifies source and destinations of Information Products.
  - *A "collection of data element instances that meet the specified requirements of a data consumer." (Wang)*

5/27/2008   **15**   Federal Railroad Administration

IP-Map — Data Attribute: Railroad Code

168

Conclusions

- *Be ready.*
  - *It should be anticipated that data quality efforts will recommend change and possibly even Change.*

- *Be clear and purposeful.*
  - *Know what you're doing and why… and be able to articulate this in practical terms that managers and workers will understand and appreciate.*

5/27/2008   **17**   Federal Railroad Administration

Questions

5/27/2008   **18**   Federal Railroad Administration

Thank You!

Scott Bernard
Chief Architect
Federal Railroad Administration
U.S. Department of Transportation
202-493-6125
scott.bernard@dot.gov

Mark Trimble
Data Architect
Federal Railroad Administration
U.S. Department of Transportation
202-493-1343
mark.trimble@dot.gov

5/27/2008   **19**   Federal Railroad Administration