Data and Information Quality in EPA's Enterprise Data Architecture

2008 Conference on Managing Environmental Quality Systems
Seattle, Washington

Briefing by:

Kevin J. Kirby, US EPA
Enterprise Data Architect

April 24, 2008

Presentation Overview

- Enterprise Architecture (EA) in Context
- Enterprise Architecture Reference Models
- The Data Management Strategy
- Addressing data quality Management
- Next Steps / Conclusion

Disclaimer:

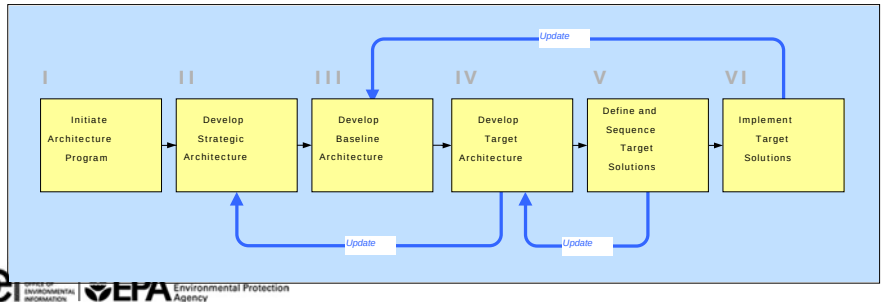
The opinions expressed in this technical presentation are those of the author and do not necessarily reflect the views of the US EPA.

1

Enterprise Architecture: What is it?

- "A strategic information asset base which defines the mission;
- the information necessary to perform the mission;
- the technologies necessary to perform the mission;
- and the transitional processes for implementing new technologies in response to changing mission needs and includes:
- a baseline architecture; a target architecture; and a sequencing plan."

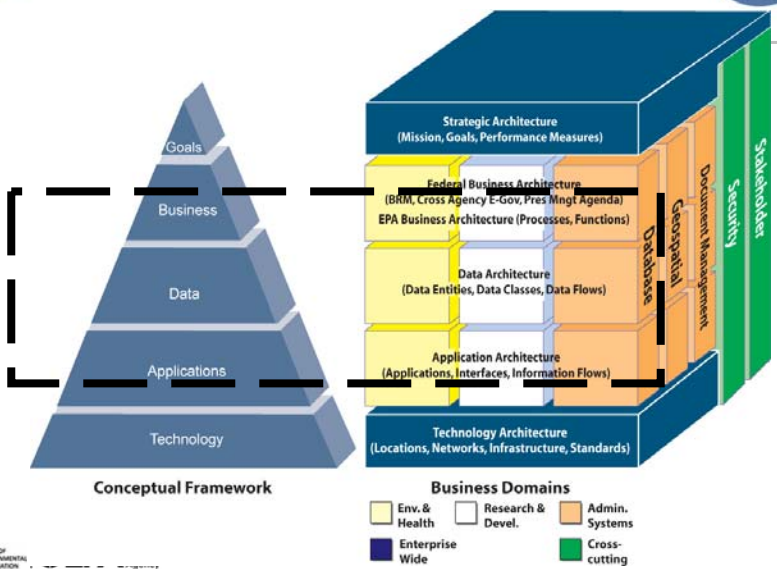
(*Source: Electronic Government Act of 2002*)



EPA Environmental Protection Agency

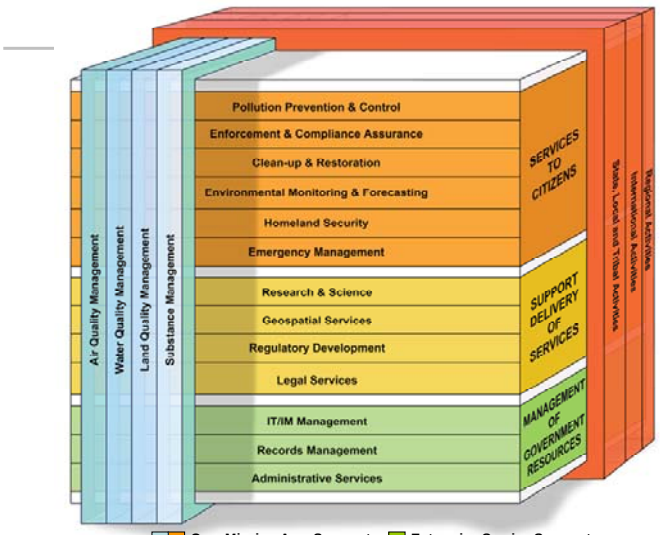
2

Enterprise Architecture Federal Conceptual Framework



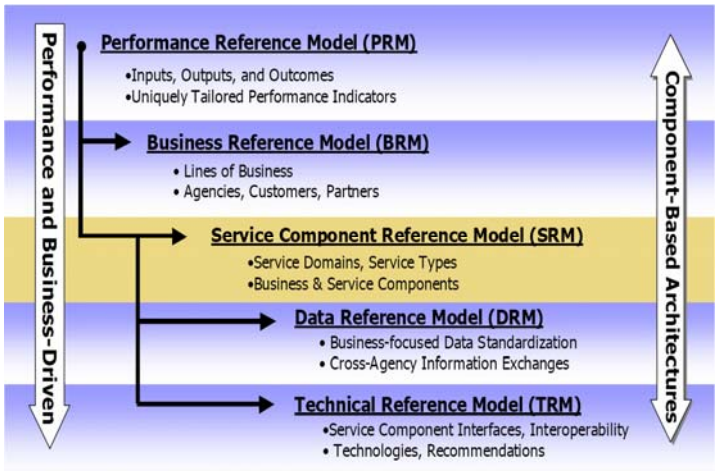
3

EPA's Enterprise Architecture Segment Model



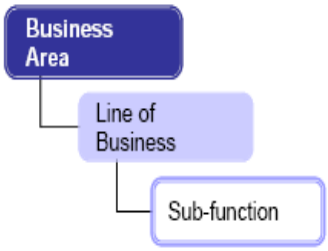
4

Federal EA Reference Models



5

FEA Business Reference Model



Busi-ness Area	LoB FEA BRM 2.0 as of May 2005	Sub-Function 1 FEA BRM 2.0 as of May 2005
Services to Citizens	108 - Environmental Management	- 023 Environmental Monitoring & Forecasting
		- 024 Environmental Remediation
		- 025 Pollution Prevention & Control

6

EPA's Business Reference Model

Business Area	LoB FEA BRM 2.0 as of May 2005	Sub-Function 1 FEA BRM 2.0 as of May 2005	Sub-Function 2 EPA BRM 4.6
Services to Citizens	108 - Environmental Management	- 023 Environmental Monitoring & Forecasting	01 Air Monitoring & Forecasting
			02 Water Monitoring & Forecasting
			03 Land Monitoring & Forecasting
		- 024 Environmental Remediation	01 Site & Area Evaluation & Cleanup
		- 025 Pollution Prevention & Control	01 Manage Air Quality
			02 Manage Water Quality
			03 Ensure Safe Drinking Water
			04 Manage Hazardous and Non-Hazardous Wastes
			05 Manage Environmental Risk from Substances
			06 Prevent Pollution
			07 Compliance & Enforcement
			08 Environmental Stewardship

Context for Data Architecture

- Most Environmental Data does not originate at EPA
- Increasing Demands for Access to Quality Information
- High Expectations to Know What Data EPA Has
- Movement across the Federal Government to Leverage Shared Services
- Federal Movement toward Outsourcing
- Alignment with the Federal Data Reference Model

OEI | EPA United States Environmental Protection Agency

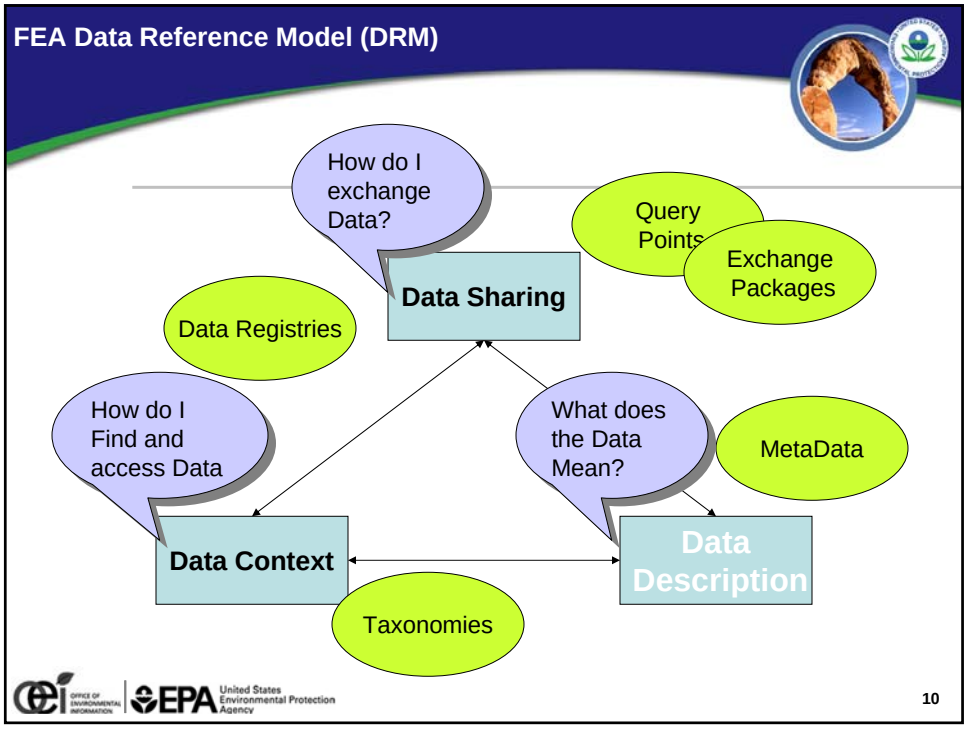
8

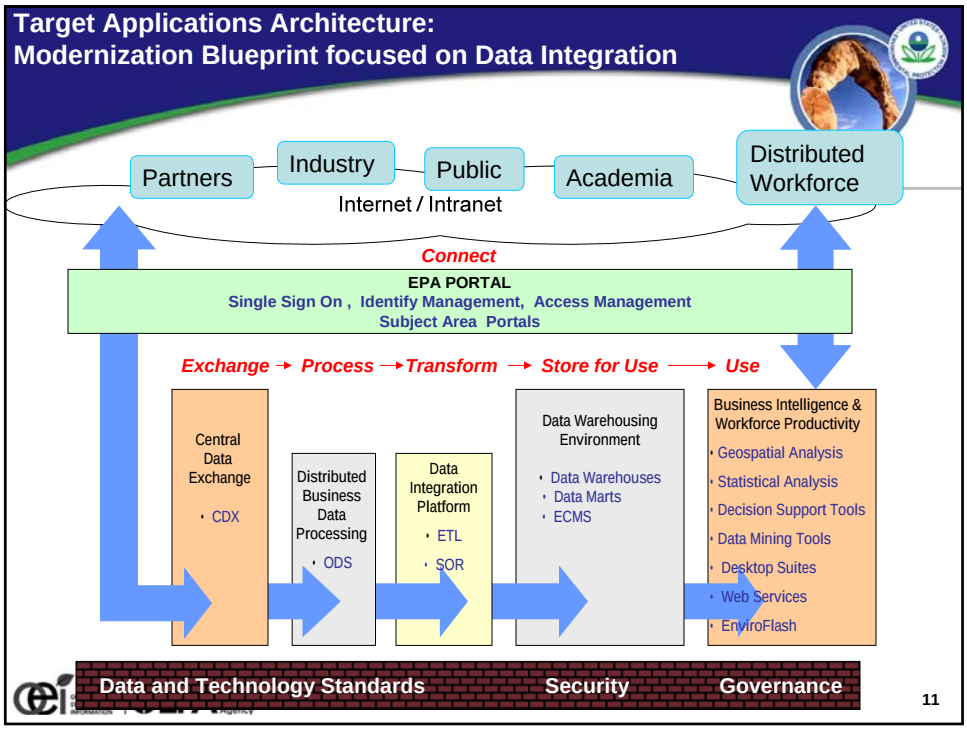
Data Architecture: Value Proposition

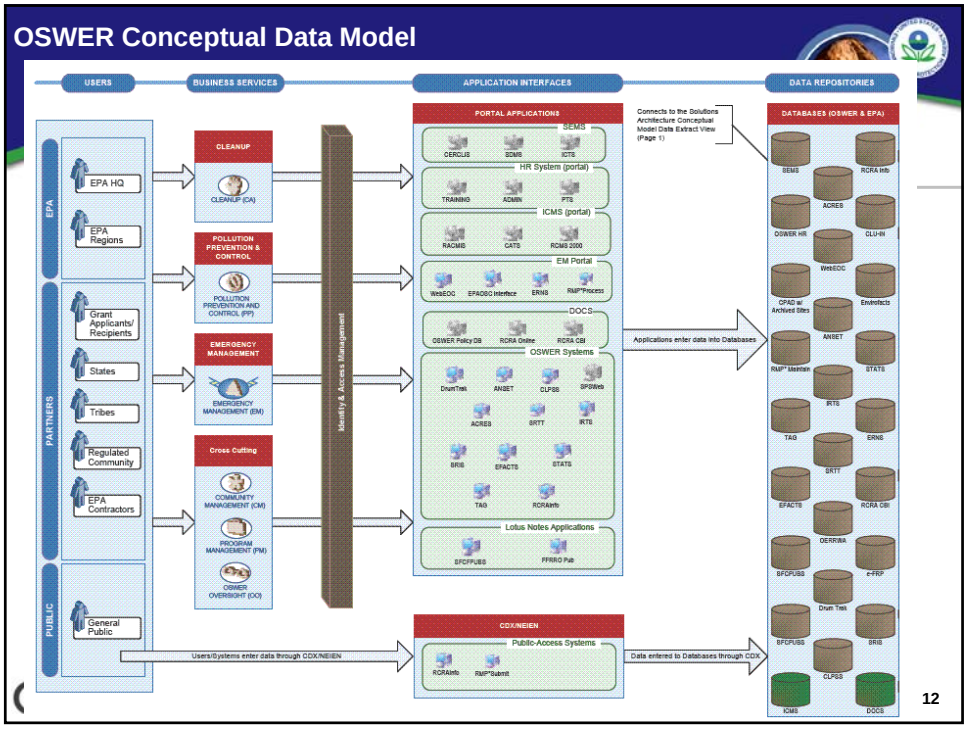
- EPA is committed to improving the understanding and use of its data assets
- To accomplish this, users and owners of data must have a common understanding of the meaning and representation of the data.
- The data must be described by documenting its characteristics or “metadata” The metadata will describe:
  - Meaning of the data (also referred to as semantic metadata)
  - Representation of the data
  - Administrative information about the data (e.g., ownership, versioning, effective dates, status)

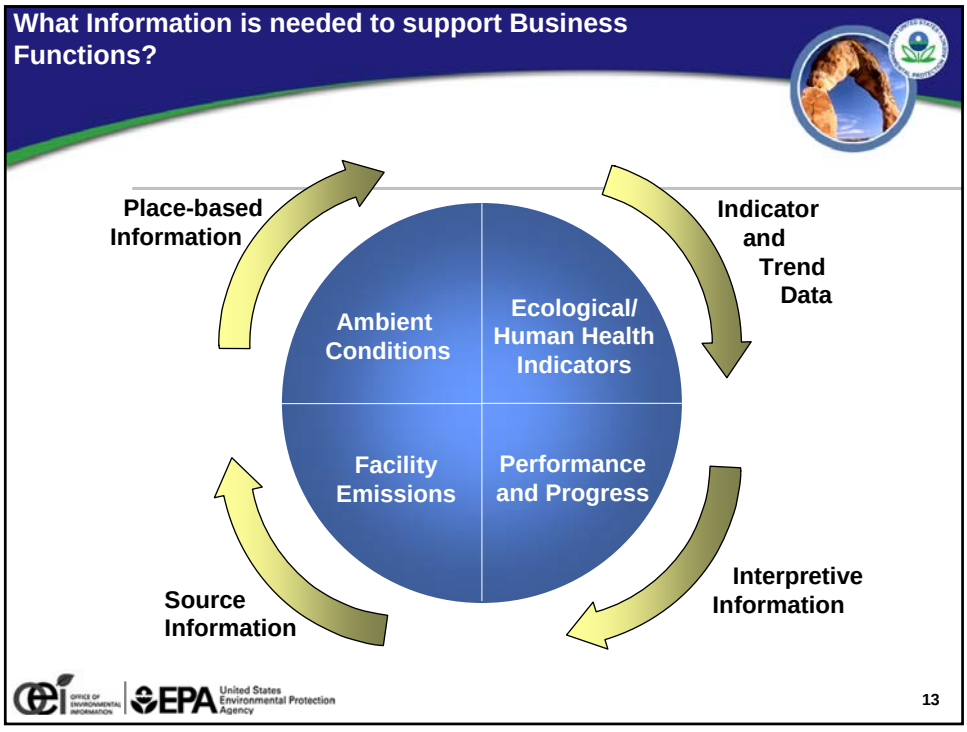
OEI | EPA United States Environmental Protection Agency

9


FEA Data Reference Model (DRM)
How do I exchange Data?
Query Points
Exchange Packages
Data Sharing
Data Registries
How do I Find and access Data
What does the Data Mean?
MetaData
Data Context
Data Description
Taxonomies
EPA United States Environmental Protection Agency
10


Target Applications Architecture:
Modernization Blueprint focused on Data Integration
Partners
Industry
Public
Academia
Distributed Workforce
Internet / Intranet
Connect
EPA PORTAL
Single Sign On , Identify Management, Access Management
Subject Area Portals
Exchange → Process → Transform → Store for Use → Use
Central Data Exchange
• CDX
Distributed Business Data Processing
• ODS
Data Integration Platform
• ETL
• SOR
Data Warehousing Environment
• Data Warehouses
• Data Marts
• ECMS
Business Intelligence & Workforce Productivity
• Geospatial Analysis
• Statistical Analysis
• Decision Support Tools
• Data Mining Tools
• Desktop Suites
• Web Services
• EnviroFlash
Data and Technology Standards
Security
Governance
11

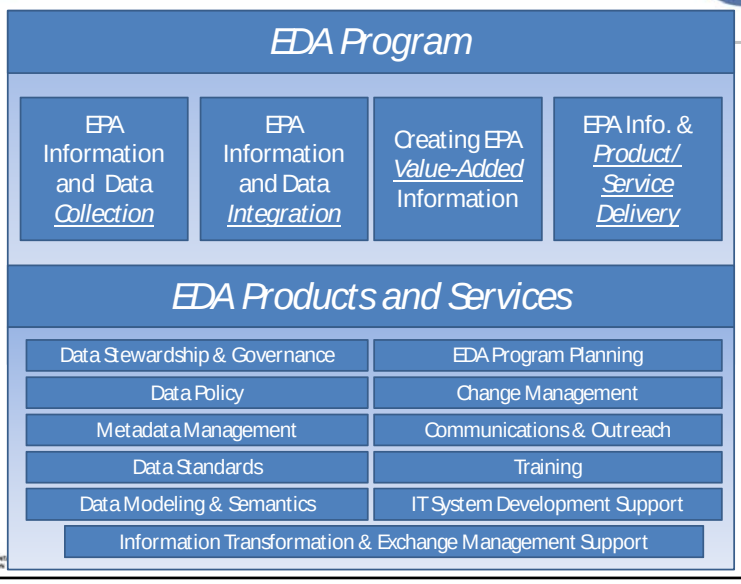

OSWER Conceptual Data Model
USERS
BUSINESS SERVICES
APPLICATION INTERFACES
DATA REPOSITORIES
EPA
PARTNERS
PUBLIC
EPA HQ
EPA Regions
Grant Applicants/ Recipients
States
Tribes
Regulated Community
EPA Contractors
General Public
CLEANUP
POLLUTION PREVENTION & CONTROL
EMERGENCY MANAGEMENT
Cross Cutting
Identity & Access Management
PORTAL APPLICATIONS
SEMS
HR System (portal)
ICMS (portal)
EM Portal
DOCS
OSWER Systems
Lotus Notes Applications
CDX/NEIEN
Public-Access Systems
Connects to the Solutions Architecture Conceptual Model Data Extract View (Page 1)
Applications enter data into Databases
Users/Systems enter data through CDX/NEIEN
Data entered to Databases through CDX
DATABASES (OSWER & EPA)
12


What Information is needed to support Business Functions?
Place-based Information
Indicator and Trend Data
Ambient Conditions
Ecological/ Human Health Indicators
Facility Emissions
Performance and Progress
Source Information
Interpretive Information
United States Environmental Protection Agency
13

The Data Management 'Wheel of Fortune'

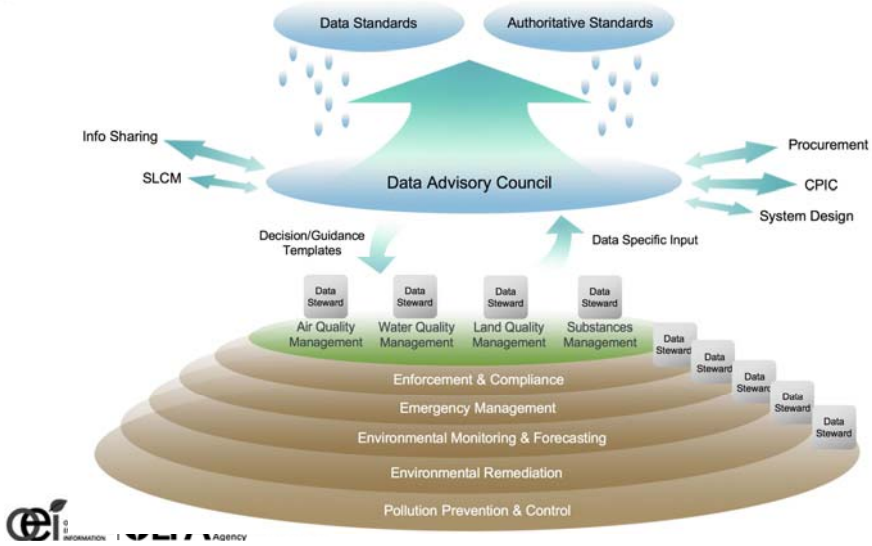
14

EPA's Enterprise Data Architecture (EDA): *Comprehensive Target Vision*



15

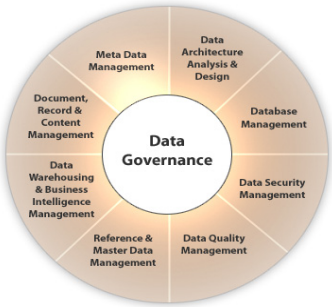
Enterprise Data Architecture: *Target Governance Model*



16

Parting Thoughts / Next Steps

- EPA has an Enterprise Architecture
- The Agency is developing a phased Data Management Framework
- Phase One includes:
  - *Enterprise Data Architecture approach*
  - *Metadata Management*
  - *Data Quality Management*
  - *Governance / Stewardship*
- Phase Two includes:
  - Data Security Management
  - *Master Data Management*
  - *Data Warehouse Management*



- The success of the EDA hinges on integrated Data Quality procedures

17

Kevin Kirby - BIO

- Kevin J. Kirby is the **Enterprise Data Architect** for the US Environmental Protection Agency, Office of Environmental Information Enterprise Architecture Team.
- His experiences in EPA have been centered in data and policy with an extensive background in the geospatial arena. Prior to joining the EPA Enterprise Architecture Team, Kevin was the IT Policy Team Lead and managed the Agency's Locational Reference Tables. He offers specific expertise in Geographic Information Systems (GIS) and watershed planning, has written the EPA's National Geospatial Data Policy, and is the Agency's representative to Federal (FGDC) National (ANSI) and International (ISO) geographic data standards organizations.
- Mr. Kirby received his academic training in Landscape Architecture, Natural Resources Planning, and a Graduate specialization in Environmental Sciences and Policy.

OEI Office of Environmental Information | EPA United States Environmental Protection Agency

18