



Curating and Cataloging Data

by Dave Wells



Module 0. About the Course (5 min)

Module 1. Self-Service Data (9 min)

- *Governance and Self-Service*
- *How We Got Here*
- *Why Self-Service Data?*
 - The World of Self-Service Data: Why?
 - Less Data Preparation Time
 - More Analysis Capacity

Module 2. Data Curation (44 min)

- *Data Curation*
 - What Does Curated Mean?
 - Focus on Data Sets
 - A Data Management Perspective
 - A Goal-Oriented Perspective
 - Practitioner Perspective
- *Why Data Curation?*
 - Purpose in the Definition
 - Big Data Influence
- *The Data Curator*
 - Curation Activities
 - Curator Qualities
 - Curators in the Organization
 - Stewards and Curators
 - Curation & Data Governance
 - Data Knowledge
- *Data Lifecycles and Curation*
 - Curating through the Lifecycle Parts 1 & 2
 - Data Lifecycles & Curation
 - Curating through the Lifecycle
- *Curating Big Data*
 - Guidelines and Practices
- *Getting Started with Data Curation*
 - The Path to Curation

Module 3. Data Cataloging (30 min)

- *Definitions*
 - Comparisons Parts 1 & 2
- *Why Data Cataloging?*
 - Analysts Working Blind
 - Improving Data Analysis
 - Supporting the Analytics Lifecycle
- *Metadata and the Catalog*
 - Scope of Metadata
 - Reading the Metadata Model
 - Data about Datasets
 - Data about Processes
 - Data for Searching
 - Data about People



DG-08: Curating and Cataloging Data

- Data about External Entities
- The Value of Metadata
- *Data Catalog Tools*
 - Catalog Functions
 - Some Examples
- *Getting Started with Data Cataloging*
 - Getting Started with Data Cataloging Parts 1 & 2

Module 4. Evaluating Data Catalog Tools (42 min)

- *The Business Case for a Data Catalog*
 - The Business Case
 - Option 1: An It Project
 - Option 2: Self-Service Without Catalog
 - Option 3: Self-Service With Catalog
 - Comparing Results
 - Annualize ROI
 - Collaborative Analytics: Parts 1 & 2
- *Kinds of Data Catalogs*
 - Variety of Tools
 - Users of Data Catalogs
- *Evaluation Criteria*
 - Cataloging Datasets
 - Cataloging Data Operations
 - Dataset Searching
 - Recommendations & Relationships
 - Data Set Evaluation
 - Data Access
 - Usage Metadata
 - Dataset Valuation
 - Metadata Catalog
 - Security
 - Data Lineage
 - Compliance
 - Data Quality
 - Data Curation
 - Socialization
 - Integration & Interoperability
 - Deployment
 - Services
 - Pricing
 - Vendor Roadmap

Module 5. Beyond Self-Service Data (44 min)

- *Information, Technology, and Data*
 - IT Divergence
 - Data Complexities
 - Changing Roles
 - New Kinds of Questions
- *The Enterprise Data Marketplace*
 - Data Services



DG-08: Curating and Cataloging Data

- A Digital Marketplace
- Marketplace Characteristics: Parts 1-3
- *EDM Architecture*
 - Marketplace Components
 - Data Marketplace Components
 - A Deeper Dive
 - Data Providers
 - Products and Inventory
 - EDM Architecture
 - The Storefront
 - The Back Room
- *EDM Data Services*
 - The Face of the Data Marketplace
 - Imagining Amazon for Data Shoppers
 - Enabling Technologies
- *Data as a Service*
 - The Cultural Shift
 - What If...Finding Data
 - What If...Data Lineage
 - What If...Getting Data
 - What If...Socializing Data
 - What If...Data Channels
 - What If...Discovering New Data
 - What If...Recommendations