

NIST Big Data Public Working Group

Overview of NIST Big Data Interoperability Framework Volume 2

Dr. Nancy Grady

Chief Data Scientist

SAIC

NIST Campus

Gaithersburg, Maryland

June 1, 2017

Presentation Overview

- Volume Presentation Outline
- Volume 1, Definitions (Nancy Grady, SAIC)
- **Volume 2, BD Taxonomies (Nancy Grady, SAIC)**
- Volume 3, Use Cases and General Requirements (Geoffrey Fox, Indiana University)
- Volume 6, Reference Architecture (David Boyd, InCadence Corp.)
- Volume 4, Security and Privacy (Arnab Roy, Fujitsu; Mark Underwood, AVP, Strategic Initiatives, Controls and Countermeasures)
- Volume 8, Reference Architecture Interface (Gregor von Laszewski, Indiana University)
- Reference Architecture Software Implementation Environment and Demonstration (Gregor von Laszewski, Indiana University)
- Volume 7, Standards Roadmap (Russell Reinsch, Center for Government Interoperability)
- Volume 9, Adoption and Modernization (Russell Reinsch, Center for Government Interoperability)

NBDIF Volume Overview

Vol. 1 BD Definitions
Defines common language

Vol. 2 BD Taxonomies
Hierarchy of NBDRA components

Vol. 3 Use Cases & Vol. 5 Arch Survey
Info gathered; requirements extracted

Vol. 6 NBDRA
Developed NBDRA

Vol.4 S&P
Interwoven topics of S&P examined

Vol. 7 Standards Roadmap
Examine standards wrt NBDRA

Vol. 8 NBDRA Interfaces
Implementation of NBDRA

Vol. 9 Adoption & Modernization



Volume Presentation Outline

- For each volume
 - Scope of the volume
 - Brief recap of version 1
 - Highlights of version 2 accomplishments
 - Summary of version 2 areas needing contributions
 - Topics that could be considered for version 3

Volume 2, Big Data Taxonomies

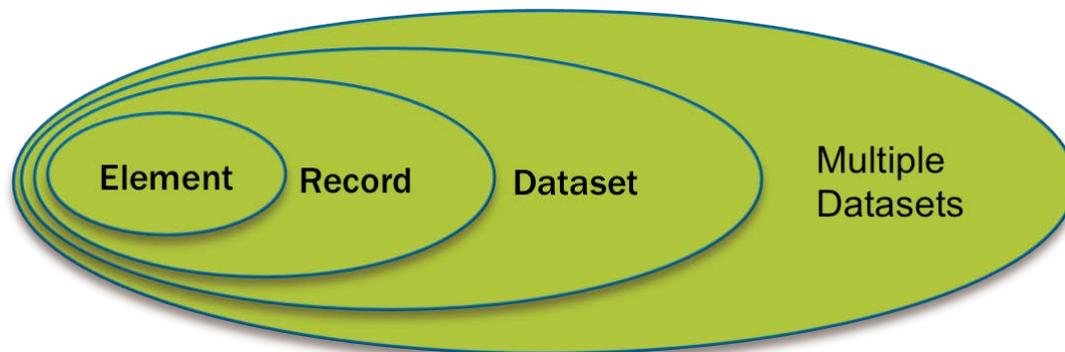
Document Scope

- Provide a taxonomy of the components of the NIST Big Data Reference Architecture (NBDRA)
 - Organize concepts
- Focus on changes resulting from the shift to the new Big Data paradigm
- Reinforce a common language
- Audience
 - New to Big Data to understand relationships of items

Volume 2, Big Data Taxonomies

Version 1 Overview

- **Collectively developed taxonomy for the NBDRA**
 - Five general architecture components and two interwoven fabrics
 - Hierarchy of actors and roles and the activities the actors perform in those roles
- **Developed a hierarchical description about data**
 - Data examined at different groupings
 - Understand what data characteristics have driven the need for the new technologies





Volume 2, Big Data Taxonomies

Version 2 Accomplishments

- Further developed the data characteristic hierarchy
- Investigated the relationship between the NBDRA taxonomy and other taxonomies

Volume 2, Big Data Taxonomies

Version 2 Opportunities for Contribution

- Suggest different organization for presentation of characteristics
 - Create actual taxonomy
- Contribute to comparison of NBDRA taxonomy to the following:
 - Cloud taxonomy
 - High-performance computing (HPC) taxonomy
 - Internet of things taxonomy
- Enhance discussion of Management fabric with respect to Big Data systems
- Discuss data virtualization with respect to multiple datasets
- Pointers to other Big Data or related taxonomies

Volume 2, Big Data Taxonomies

Possible Version 3 Topics

- Comparison with other Big Data Taxonomies
- Features taxonomy of storage techniques
- Comparison with additional related Taxonomies
- Pointers to taxonomies for
 - Data Types
 - Analytics
 - Visualization
- Creation of data characteristic Taxonomy

Breakout 1, Definitions and Taxonomies, Breakout Plan

- Review Version 2 slide of remaining items
 - Do any need so much work they should be deferred to version 3
- What have we missed
- What is not needed or is poorly expressed
- Review of Version 3 slide
 - Anything that should be put in version 2
 - Anything missing