

NIST Special Publication 1500-7

**NIST Big Data Interoperability
Framework:
Volume 7, Standards Roadmap**

Final Version 1

NIST Big Data Public Working Group
Technology Roadmap Subgroup

This publication is available free of charge from:
<http://dx.doi.org/10.6028/NIST.SP.1500-7>

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

NIST Special Publication 1500-7

NIST Big Data Interoperability Framework: Volume 7, Standards Roadmap

Final Version 1

NIST Big Data Public Working Group (NBD-PWG)
Technology Roadmap Subgroup
Information Technology Laboratory

This publication is available free of charge from:
<http://dx.doi.org/10.6028/NIST.SP.1500-7>

September 2015



U. S. Department of Commerce
Penny Pritzker, Secretary

National Institute of Standards and Technology
Willie May, Under Secretary of Commerce for Standards and Technology and Director

National Institute of Standards and Technology (NIST) Special Publication 1500-7
46 pages (September 16, 2015)

NIST Special Publication series 1500 is intended to capture external perspectives related to NIST standards, measurement, and testing-related efforts. These external perspectives can come from industry, academia, government, and others. These reports are intended to document external perspectives and do not represent official NIST positions.

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

There may be references in this publication to other publications currently under development by NIST in accordance with its assigned statutory responsibilities. The information in this publication, including concepts and methodologies, may be used by federal agencies even before the completion of such companion publications. Thus, until each publication is completed, current requirements, guidelines, and procedures, where they exist, remain operative. For planning and transition purposes, federal agencies may wish to closely follow the development of these new publications by NIST.

Organizations are encouraged to review all draft publications during public comment periods and provide feedback to NIST. All NIST publications are available at <http://www.nist.gov/publication-portal.cfm>.

National Institute of Standards and Technology
Attn: Wo Chang, Information Technology Laboratory
100 Bureau Drive (Mail Stop 8900) Gaithersburg, MD 20899-8930
Email: SP1500comments@nist.gov

Reports on Computer Systems Technology

The Information Technology Laboratory (ITL) at NIST promotes the U.S. economy and public welfare by providing technical leadership for the Nation's measurement and standards infrastructure. ITL develops tests, test methods, reference data, proof of concept implementations, and technical analyses to advance the development and productive use of information technology (IT). ITL's responsibilities include the development of management, administrative, technical, and physical standards and guidelines for the cost-effective security and privacy of other than national security-related information in federal information systems. This document reports on ITL's research, guidance, and outreach efforts in IT and its collaborative activities with industry, government, and academic organizations.

Abstract

Big Data is a term used to describe the large amount of data in the networked, digitized, sensor-laden, information-driven world. While opportunities exist with Big Data, the data can overwhelm traditional technical approaches, and the growth of data is outpacing scientific and technological advances in data analytics. To advance progress in Big Data, the NIST Big Data Public Working Group (NBD-PWG) is working to develop consensus on important fundamental concepts related to Big Data. The results are reported in the *NIST Big Data Interoperability Framework* series of volumes. This volume, Volume 7, contains summaries of the work presented in the other six volumes and an investigation of standards related to Big Data.

Keywords

Big Data; Big Data Application Provider; Big Data characteristics; Big Data taxonomy; Big Data standards; Data Consumer; Data Provider; Management Fabric; reference architecture; Security and Privacy Fabric; System Orchestrator; use cases.

Acknowledgements

This document reflects the contributions and discussions by the membership of the NBD-PWG, co-chaired by Wo Chang of the NIST IITL, Robert Marcus of ET-Strategies, and Chaitanya Baru, University of California San Diego Supercomputer Center.

The document contains input from members of the NBD-PWG: Technology Roadmap Subgroup, led by Carl Buffington (Vistrionix), David Boyd (InCadence Strategic Solutions), and Dan McClary (Oracle); Definitions and Taxonomies Subgroup, led by Nancy Grady (SAIC), Natasha Balac (SDSC), and Eugene Luster (R2AD); Use Cases and Requirements Subgroup, led by Geoffrey Fox (University of Indiana) and Tsegereda Beyene (Cisco); Security and Privacy Subgroup, led by Arnab Roy (Fujitsu) and Akhil Manchanda (GE); and Reference Architecture Subgroup, led by Orit Levin (Microsoft), Don Krapohl (Augmented Intelligence), and James Ketner (AT&T).

NIST SP1500-7, Version 1, has been collaboratively authored by the NBD-PWG. As of the date of this publication, there are over six hundred NBD-PWG participants from industry, academia, and government. Federal agency participants include the National Archives and Records Administration (NARA), National Aeronautics and Space Administration (NASA), National Science Foundation (NSF), and the U.S. Departments of Agriculture, Commerce, Defense, Energy, Health and Human Services, Homeland Security, Transportation, Treasury, and Veterans Affairs.

NIST acknowledges the specific contributions^a to this volume by the following NBD-PWG members:

Chaitan Baru <i>University of California, San Diego, Supercomputer Center</i>	Pavithra Kenjige <i>PK Technologies</i>	Shawn Miller <i>U.S. Department of Veterans Affairs</i>
David Boyd <i>InCadence Strategic Services</i>	Brenda Kirkpatrick <i>Hewlett-Packard</i>	William Miller <i>MaCT USA</i>
Carl Buffington <i>Vistrionix</i>	Donald Krapohl <i>Augmented Intelligence</i>	Sanjay Mishra <i>Verizon</i>
Pw Carey <i>Compliance Partners, LLC</i>	Luca Lepori <i>Data Hold</i>	Quyen Nguyen <i>NARA</i>
Wo Chang <i>NIST</i>	Orit Levin <i>Microsoft</i>	John Rogers <i>Hewlett-Packard</i>
Yuri Demchenko <i>University of Amsterdam</i>	Jan Levine <i>kloudtrack</i>	Doug Scrimager <i>Slalom Consulting</i>
Nancy Grady <i>SAIC</i>	Serge Mankovski <i>CA Technologies</i>	Cherry Tom <i>IEEE-SA</i>
Keith Hare <i>JCC Consulting, Inc.</i>	Robert Marcus <i>ET-Strategies</i>	Timothy Zimmerlin <i>Automation Technologies Inc.</i>
Bruno Kelpsas <i>Microsoft Consultant</i>	Gary Mazzaferro <i>AlloyCloud, Inc.</i>	

The editors for this document were David Boyd, Carl Buffington, and Wo Chang

^a “Contributors” are members of the NIST Big Data Public Working Group who dedicated great effort to prepare and substantial time on a regular basis to research and development in support of this document.

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDFP	S&P	M
W3C HTML5 A vocabulary and associated APIs for HTML and XHTML	This specification defines the 5th major revision of the core language of the World Wide Web—HTML.		I	U	I/U			
W3C Internationalization Tag Set (ITS) 2.0	The ITS 2.0 specification enhances the foundation to integrate automated processing of human language into core Web technologies and concepts that are designed to foster the automated creation and processing of multilingual Web content.		I	U	I/U	I/U		
W3C OWL 2 Web Ontology Language	The OWL 2 Web Ontology Language, informally OWL 2, is an ontology language for the Semantic Web with formally defined meaning.		I	U	I/U	I/U		
W3C Platform for Privacy Preferences (P3P) 1.0	The P3P enables Web sites to express their privacy practices in a standard format that can be retrieved automatically and interpreted easily by user agents.		I	U	I/U		I/U	
W3C Protocol for Web Description Resources (POWDER)	POWDER—the Protocol for Web Description Resources—provides a mechanism to describe and discover Web resources and helps the users to make a decision whether a given resource is of interest.		I	U	I/U			
W3C Provenance	Provenance is information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability, or trustworthiness. The Provenance Family of Documents (PROV) defines a model, corresponding serializations, and other supporting definitions to enable the interoperable interchange of provenance information in heterogeneous environments such as the Web.		I	U	I/U	I/U	U	

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDF P	S&P	M
W3C Rule Interchange Format (RIF)	RIF is a series of standards for exchanging rules among rule systems, in particular among Web rule engines.		I	U	I/U	I/U		
W3C Service Modeling Language (SML) 1.1	This specification defines the SML, Version 1.1 used to model complex services and systems, including their structure, constraints, policies, and best practices.	I/U	I	U	I/U			
W3C Simple Knowledge Organization System Reference (SKOS)	This document defines the SKOS, a common data model for sharing and linking knowledge organization systems via the Web.		I	U	I/U			
W3C Simple Object Access Protocol (SOAP) 1.2	SOAP is a protocol specification for exchanging structured information in the implementation of web services in computer networks.		I	U	I/U			
W3C SPARQL 1.1	SPARQL is a language specification for the query and manipulation of linked data in a RDF format.		I	U	I/U	I/U		
W3C Web Service Description Language (WSDL) 2.0	This specification describes the WSDL Version 2.0, an XML language for describing Web services.	U	I	U	I/U			
W3C XML Key Management Specification (XKMS) 2.0	This standard specifies protocols for distributing and registering public keys, suitable for use in conjunction with the W3C Recommendations for XML Signature [XML-SIG] and XML Encryption [XML-Enc]. The XKMS comprises two parts — the XML Key Information Service Specification (X-KISS) and the XML Key Registration Service Specification (X-KRSS).	U	I	U	I/U			
OGC® OpenGIS® Catalogue Services Specification 2.0.2 - ISO Metadata Application Profile	This series of standard covers Catalogue Services based on ISO19115/ISO19119, which are organized and implemented for the discovery, retrieval and management of data metadata, services metadata, and application metadata.		I	U	I/U			

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDF P	S&P	M
OGC® OpenGIS® GeoAPI	The GeoAPI Standard defines, through the GeoAPI library, a Java language API including a set of types and methods which can be used for the manipulation of geographic information structured following the specifications adopted by the Technical Committee 211 of the ISO and by the OGC®.		I	U	I/U	I/U		
OGC® OpenGIS® GeoSPARQL	The OGC® GeoSPARQL standard supports representing and querying geospatial data on the Semantic Web. GeoSPARQL defines a vocabulary for representing geospatial data in RDF, and it defines an extension to the SPARQL query language for processing geospatial data.		I	U	I/U	I/U		
OGC® OpenGIS® Geography Markup Language (GML) Encoding Standard	The GML is an XML grammar for expressing geographical features. GML serves as a modeling language for geographic systems as well as an open interchange format for geographic transactions on the Internet.		I	U	I/U	I/U		
OGC® Geospatial eXtensible Access Control Markup Language (GeoXACML) Version 1	The Policy Language introduced in this document defines a geo-specific extension to the XACML Policy Language, as defined by the OASIS standard eXtensible Access Control Markup Language (XACML), Version 2.0.		I	U	I/U	I/U	I/U	
OGC® Network Common Data Form (netCDF)	netCDF is a set of software libraries and self-describing, machine-independent data formats that support the creation, access, and sharing of array-oriented scientific data.		I	U	I/U			

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDF P	S&P	M
OGC® Open Modelling Interface Standard (OpenMI)	The purpose of the OpenMI is to enable the runtime exchange of data between process simulation models and also between models and other modelling tools such as databases and analytical and visualization applications.		I	U	I/U	I/U		
OGC® OpenSearch Geo and Time Extensions	This OGC standard specifies the Geo and Time extensions to the OpenSearch query protocol. OpenSearch is a collection of simple formats for the sharing of search results.		I	U	I/U	I		
OGC® Web Services Context Document (OWS Context)	The OGC® OWS Context was created to allow a set of configured information resources (service set) to be passed between applications primarily as a collection of services.		I	U	I/U	I		
OGC® Sensor Web Enablement (SWE)	This series of standards support interoperability interfaces and metadata encodings that enable real-time integration of heterogeneous sensor webs. These standards include a modeling language (SensorML), common data model, and sensor observation, planning, and alerting service interfaces.		I	U	I/U			
OGC® OpenGIS® Simple Features Access	Describes the common architecture for simple feature geometry and is also referenced as ISO 19125. It also implements a profile of the spatial schema described in ISO 19107:2003.		I	U	I/U	I/U		
OGC® OpenGIS® Georeferenced Table Joining Service (TJS) Implementation Standard	This standard is the specification for a TJS that defines a simple way to describe and exchange tabular data that contains information about geographic objects.		I	U	I/U	I/U		

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDF P	S&P	M
OGC® OpenGIS® Web Coverage Processing Service Interface (WCPS) Standard	Defines a protocol-independent language for the extraction, processing, and analysis of multi-dimensional gridded coverages representing sensor, image, or statistics data.		I	U	I/U	I		
OGC® OpenGIS® Web Coverage Service (WCS)	This document specifies how a WCS offers multidimensional coverage data for access over the Internet. This document specifies a core set of requirements that a WCS implementation must fulfill.		I	U	I/U	I		
OGC® Web Feature Service (WFS) 2.0 Interface Standard	The WFS standard provides for fine-grained access to geographic information at the feature and feature property level. This International Standard specifies discovery operations, query operations, locking operations, transaction operations, and operations to manage stored, parameterized query expressions.		I	U	I/U	I		
OGC® OpenGIS® Web Map Service (WMS) Interface Standard	The OpenGIS® WMS Interface Standard provides a simple HTTP interface for requesting geo-registered map images from one or more distributed geospatial databases.		I	U	I/U	I		
OGC® OpenGIS® Web Processing Service (WPS) Interface Standard	The OpenGIS® WPS Interface Standard provides rules for standardizing how inputs and outputs (requests and responses) for geospatial processing services, such as polygon overlay. The standard also defines how a client can request the execution of a process, and how the output from the process is handled. It defines an interface that facilitates the publishing of geospatial processes and clients' discovery of and binding to those processes.		I	U	I/U	I		
OASIS AS4 Profile of ebMS 3.0 v1.0	Standard for business to business exchange of messages via a web service platform.		I	U	I/U			

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDF P	S&P	M
OASIS Advanced Message Queuing Protocol (AMQP) Version 1.0	The AMQP is an open internet protocol for business messaging. It defines a binary wire-level protocol that allows for the reliable exchange of business messages between two parties.		I	U	U	I		
OASIS Application Vulnerability Description Language (AVDL) v1.0	This specification describes a standard XML format that allows entities (such as applications, organizations, or institutes) to communicate information regarding web application vulnerabilities.		I	U	I		U	
OASIS Biometric Identity Assurance Services (BIAS) Simple Object Access Protocol (SOAP) Profile v1.0	This OASIS BIAS profile specifies how to use XML (XML10) defined in ANSI INCITS 442-2010—BIAS to invoke SOAP-based services that implement BIAS operations.		I	U	I/U		U	
OASIS Content Management Interoperability Services (CMIS)	The CMIS standard defines a domain model and set of bindings that include Web Services and ReSTful AtomPub that can be used by applications to work with one or more Content Management repositories/systems.		I	U	I/U	I		
OASIS Digital Signature Service (DSS)	This specification describes two XML-based request/response protocols - a signing protocol and a verifying protocol. Through these protocols, a client can send documents (or document hashes) to a server and receive back a signature on the documents; or send documents (or document hashes) and a signature to a server, and receive back an answer on whether the signature verifies the documents.		I	U	I/U			

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDFP	S&P	M
OASIS Directory Services Markup Language (DSML) v2.0	The DSML provides a means for representing directory structural information as an XML document methods for expressing directory queries and updates (and the results of these operations) as XML documents		I	U	I/U	I		
OASIS ebXML Messaging Services	These specifications define a communications-protocol neutral method for exchanging electronic business messages as XML.		I	U	I/U			
OASIS ebXML RegRep	ebXML RegRep is a standard defining the service interfaces, protocols, and information model for an integrated registry and repository. The repository stores digital content while the registry stores metadata that describes the content in the repository.		I	U	I/U	I		
OASIS ebXML Registry Information Model	The Registry Information Model provides a blueprint or high-level schema for the ebXML Registry. It provides implementers with information on the type of metadata that is stored in the Registry as well as the relationships among metadata Classes.		I	U	I/U			
OASIS ebXML Registry Services Specification	An ebXML Registry is an information system that securely manages any content type and the standardized metadata that describes it. The ebXML Registry provides a set of services that enable sharing of content and metadata between organizational entities in a federated environment.		I	U	I/U			
OASIS eXtensible Access Control Markup Language (XACML)	The standard defines a declarative access control policy language implemented in XML and a processing model describing how to evaluate access requests according to the rules defined in policies.		I	U	I/U	I/U	I/U	

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDFP	S&P	M
OASIS Message Queuing Telemetry Transport (MQTT)	MQTT is a Client Server publish/subscribe messaging transport protocol for constrained environments such as for communication in Machine to Machine and Internet of Things contexts where a small code footprint is required and/or network bandwidth is at a premium.		I	U	I/U			
OASIS Open Data (OData) Protocol	The OData Protocol is an application-level protocol for interacting with data via RESTful interfaces. The protocol supports the description of data models and the editing and querying of data according to those models.		I	U	I/U	I/U		
OASIS Search Web Services (SWS)	The OASIS SWS initiative defines a generic protocol for the interaction required between a client and server for performing searches. SWS define an Abstract Protocol Definition to describe this interaction.		I	U	I/U			
OASIS Security Assertion Markup Language (SAML) v2.0	The SAML defines the syntax and processing semantics of assertions made about a subject by a system entity. This specification defines both the structure of SAML assertions and an associated set of protocols, in addition to the processing rules involved in managing a SAML system.		I	U	I/U	I/U	I/U	
OASIS SOAP-over-UDP (User Datagram Protocol) v1.1	This specification defines a binding of SOAP to user datagrams, including message patterns, addressing requirements, and security considerations.		I	U	I/U			

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDF P	S&P	M
OASIS Solution Deployment Descriptor Specification v1.0	This specification defines schema for two XML document types: Package Descriptors and Deployment Descriptors. Package Descriptors define characteristics of a package used to deploy a solution. Deployment Descriptors define characteristics of the content of a solution package, including the requirements that are relevant for creation, configuration, and maintenance of the solution content.	U						I/U
OASIS Symptoms Automation Framework (SAF) Version 1.0	This standard defines reference architecture for the Symptoms Automation Framework, a tool in the automatic detection, optimization, and remediation of operational aspects of complex systems.							I/U
OASIS Topology and Orchestration Specification for Cloud Applications Version 1.0	The concept of a “service template” is used to specify the “topology” (or structure) and “orchestration” (or invocation of management behavior) of IT services. This specification introduces the formal description of Service Templates, including their structure, properties, and behavior.	I/U			U	I		I/U
OASIS Universal Business Language (UBL) v2.1	The OASIS UBL defines a generic XML interchange format for business documents that can be restricted or extended to meet the requirements of particular industries.		I	U	I/U	U		
OASIS Universal Description, Discovery and Integration (UDDI) v3.0.2	The focus of UDDI is the definition of a set of services supporting the description and discovery of (1) businesses, organizations, and other Web services providers, (2) the Web services they make available, and (3) the technical interfaces which may be used to access those services.		I	U	I/U			U

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDF P	S&P	M
OASIS Unstructured Information Management Architecture (UIMA) v1.0	The UIMA specification defines platform-independent data representations and interfaces for text and multimodal analytics.				U	I		
OASIS Unstructured Operation Markup Language (UOML) v1.0	UOML is an interface standard to process unstructured documents; it plays the similar role as SQL to structured data. UOML is expressed with standard XML.		I	U	I/U	I		
OASIS/W3C WebCGM v2.1	Computer Graphics Metafile (CGM) is an ISO standard, defined by ISO/IEC 8632:1999, for the interchange of 2D vector and mixed vector/raster graphics. WebCGM is a profile of CGM, which adds Web linking and is optimized for Web applications in technical illustration, electronic documentation, geophysical data visualization, and similar fields.		I	U	I/U	I		
OASIS Web Services Business Process Execution Language (WS-BPEL) v2.0	This standard defines a language for specifying business process behavior based on Web services. WS-BPEL provides a language for the specification of Executable and Abstract business processes.	U			I			
OASIS/W3C - Web Services Distributed Management (WSDM): Management Using Web Services (MUWS) v1.1	MUWS defines how an IT resource connected to a network provides manageability interfaces such that the IT resource can be managed locally and from remote locations using Web services technologies.	U			I	I	U	U
OASIS WSDM: Management of Web Services (MOWS) v1.1	This part of the WSDM specification addresses management of the Web services endpoints using Web services protocols.	U			I	I	U	U
OASIS Web Services Dynamic Discovery (WS-Discovery) v1.1	This specification defines a discovery protocol to locate services. The primary scenario for discovery is a client searching for one or more target services.	U	I	U	I/U			U

Standard Name/Number	Description	NBDRA Components						
		SO	DP	DC	BDAP	BDFP	S&P	M
OASIS Web Services Federation Language (WS-Federation) v1.2	This specification defines mechanisms to allow different security realms to federate, such that authorized access to resources managed in one realm can be provided to security principals whose identities and attributes are managed in other realms.		I	U	I/U		U	
OASIS Web Services Notification (WSN) v1.3	WSN is a family of related specifications that define a standard Web services approach to notification using a topic-based publish/subscribe pattern.		I	U	I/U			
IETF Simple Network Management Protocol (SNMP) v3	SNMP is a series of IETF-sponsored standards for remote management of system/network resources and transmission of status regarding network resources. The standards include definitions of standard management objects along with security controls.				I	I	I/U	U
IETF Extensible Provisioning Protocol (EPP)	This IETF series of standards describes an application-layer client-server protocol for the provisioning and management of objects stored in a shared central repository. Specified in XML, the protocol defines generic object management operations and an extensible framework that maps protocol operations to objects.	U						I/U

Table Notes:

- SO = System Orchestrator component
- DP = Data Provider component
- DC = Data Consumer component
- BDAP = Big Data Application Provider component
- BDFP = Big Data Framework Provider component
- S&P = Security and Privacy Fabric
- M = Management Fabric

6.2 GAP IN STANDARDS

The potential gaps in Big Data standardization are provided in this section to describe broad areas that may be of interest to SDOs, consortia, and readers of this document. The list provided below was produced by an ISO/IEC Joint Technical Committee 1 (JTC1) Study Group on Big Data to serve as a potential guide to ISO in their establishment of Big Data standards activities.³ The potential Big Data standardization gaps, identified by the study group, described broad areas that may be of interest to this community. These gaps in standardization activities related to Big Data are in the following areas:

1. Big Data use cases, definitions, vocabulary and reference architectures (e.g., system, data, platforms, online/offline)
2. Specifications and standardization of metadata including data provenance
3. Application models (e.g., batch, streaming)
4. Query languages including non-relational queries to support diverse data types (e.g., XML, RDF, JSON, multimedia) and Big Data operations (e.g., matrix operations)
5. Domain-specific languages
6. Semantics of eventual consistency
7. Advanced network protocols for efficient data transfer
8. General and domain-specific ontologies and taxonomies for describing data semantics including interoperation between ontologies
9. Big Data security and privacy access controls
10. Remote, distributed, and federated analytics (taking the analytics to the data) including data and processing resource discovery and data mining
11. Data sharing and exchange
12. Data storage (e.g., memory storage system, distributed file system, data warehouse)
13. Human consumption of the results of big data analysis (e.g., visualization)
14. Energy measurement for Big Data
15. Interface between relational (i.e., SQL) and non-relational (i.e., not only [or no] Structured Query Language [NoSQL]) data stores
16. Big Data quality and veracity description and management

6.3 PATHWAY TO ADDRESS STANDARDS GAPS

Standards often evolve from implementation of best practices and approaches which are proven against real-world applications or from theory that is tuned to reflect additional variables and conditions uncovered during implementation. In the case of Big Data, most standards are evolving from existing standards modified to address the unique characteristics of Big Data. Like many terms that have come into common usage in the current information age, Big Data has many possible meanings depending on the context from which it is viewed. Big Data discussions are complicated by the lack of accepted definitions, taxonomies, and common reference views. The products of the NBD-PWG are designed to specifically address the lack of consistency. Recognizing this lack of a common framework on which to build standards, ISO/IEC JTC1 has specifically chartered a working group, which will first focus on developing common definitions and a reference architecture. Once established, the definitions and reference architecture will form the basis for evolution of existing standards to meet the unique needs of Big Data and evaluation of existing implementations and practices as candidates for new Big Data-related standards. In the first case, existing standards efforts may address these gaps by either expanding or adding to the existing standard to accommodate Big Data characteristics or developing Big Data unique profiles within the framework of the existing standards. The exponential growth of data is already resulting in the development of new theories addressing topics from synchronization of data across large distributed computing environments to addressing consistency in high volume and velocity environments. As actual implementations of technologies are proven, reference implementations will evolve based on community-accepted open source efforts.

Acronyms A: Acronyms

AMQP	Advanced Message Queuing Protocol
ANSI	American National Standards Institute
API	application programming interface
AVDL	Application Vulnerability Description Language
BDAP	Big Data Application Provider component
BDFP	Big Data Framework Provider component
BIAS	Biometric Identity Assurance Services
CGM	Computer Graphics Metafile
CIA	confidentiality, integrity, and availability
CMIS	Content Management Interoperability Services
CPR	Capability Provider Requirements
DC	Data Consumer component
DCAT	Data Catalog Vocabulary
DCR	Data Consumer Requirements
DOM	Document Object Model
DP	Data Provider component
DSML	Directory Services Markup Language
DSR	Data Source Requirements
DSS	Digital Signature Service
EPP	Extensible Provisioning Protocol
EXI	Efficient XML Interchange
GeoXACML	Geospatial eXtensible Access Control Markup Language
GML	Geography Markup Language
GRC	governance, risk management, and compliance
HTML	HyperText Markup Language
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IETF	Internet Engineering Task Force
INCITS	International Committee for Information Technology Standards
ISO	International Organization for Standardization
IT	information technology
ITL	Information Technology Laboratory
ITS	Internationalization Tag Set
JPEG	Joint Photographic Experts Group
JSON	JavaScript Object Notation
JSR	Java Specification Request
JTC1	Joint Technical Committee 1
LMR	Life cycle Management Requirements
M	Management Fabric
MFI	Metamodel Framework for Interoperability
MOWS	Management of Web Services
MPEG	Moving Picture Experts Group
MQTT	Message Queuing Telemetry Transport
MUWS	Management Using Web Services
NARA	National Archives and Records Administration
NASA	National Aeronautics and Space Administration
NBD-PWG	NIST Big Data Public Working Group
NCAP	Network Capable Application Processor

netCDF	Network Common Data Form
NIST	National Institute of Standards and Technology
NoSQL	Not only (or no) Structured Query Language
NSF	National Science Foundation
OASIS	Organization for the Advancement of Structured Information Standards
OData	Open Data
ODMS	On-Demand Model Selection
OGC	Open Geospatial Consortium
OpenMI	Open Modelling Interface Standard
OR	Other Requirements
OWS Context	OGC Web Services Context Document
P3P	Platform for Privacy Preferences Project
PICS	Platform for Internet Content Selection
POWDER	Protocol for Web Description Resources
RDF	Resource Description Framework
RFID	radio frequency identification
RIF	Rule Interchange Format
S&P	Security and Privacy Fabric
SAF	Symptoms Automation Framework
SAML	Security Assertion Markup Language
SDO	standards development organization
SKOS	Simple Knowledge Organization System Reference
SLA	service-level agreement
SML	Service Modeling Language
SNMP	Simple Network Management Protocol
SO	System Orchestrator component
SOAP	Simple Object Access Protocol
SPR	Security and Privacy Requirements
SQL	Structured Query Language
SWE	Sensor Web Enablement
SWS	Search Web Services
TEDS	Transducer Electronic Data Sheet
TJS	Table Joining Service
TPR	Transformation Provider Requirements
TR	Technical Report
UBL	Universal Business Language
UDDI	Universal Description, Discovery, and Integration
UDP	User Datagram Protocol
UIMA	Unstructured Information Management Architecture
UOML	Unstructured Operation Markup Language
W3C	World Wide Web Consortium
WCPS	Web Coverage Processing Service Interface
WCS	Web Coverage Service
WFS	Web Feature Service
WMS	Web Map Service
WPS	Web Processing Service
WS-BPEL	Web Services Business Process Execution Language
WS-Discovery	Web Services Dynamic Discovery
WSDL	Web Services Description Language
WSDM	Web Services Distributed Management
WS-Federation	Web Services Federation Language

WSN	Web Services Notification
XACML	eXtensible Access Control Markup Language
XDM	XPath Data Model
X-KISS	XML Key Information Service Specification
XKMS	XML Key Management Specification
X-KRSS	XML Key Registration Service Specification
XMI	XML Metadata Interchange
XML	Extensible Markup Language
XSLT	XSL Transformations

Appendix B: References

GENERAL RESOURCES

Institute of Electrical and Electronics Engineers (IEEE). <https://www.ieee.org/index.html>

International Committee for Information Technology Standards (INCITS). <http://www.incits.org/>

International Electrotechnical Commission (IEC). <http://www.iec.ch/>

International Organization for Standardization (ISO). <http://www.iso.org/iso/home.html>

Open Geospatial Consortium (OGC). <http://www.opengeospatial.org/>

Open Grid Forum (OGF). <https://www.ogf.org/ogf/doku.php>

Organization for the Advancement of Structured Information Standards (OASIS). <https://www.oasis-open.org/>

World Wide Web Consortium (W3C). <http://www.w3.org/>

DOCUMENT REFERENCES

¹ The White House Office of Science and Technology Policy, “Big Data is a Big Deal,” *OSTP Blog*, accessed February 21, 2014, <http://www.whitehouse.gov/blog/2012/03/29/big-data-big-deal>.

² Cloud Security Alliance, *Expanded Top Ten Big Data Security and Privacy Challenges*, April 2013. https://downloads.cloudsecurityalliance.org/initiatives/bdwg/Expanded_Top_Ten_Big_Data_Security_and_Privacy_Challenges.pdf.

³ “Big Data, Preliminary Report 2014,” ISO/IEC JTC1: Information Technology. http://www.iso.org/iso/big_data_report-jtc1.pdf. (Accessed March 2, 2015). Pages 21-23.