NIST Big Data Public Working Group (NBD-PWG)/Subgroups Joint Meeting

Wo Chang, NIST
Robert Marcus, ET-Strategies
Chaitanya Baru, UC San Diego
http://bigdatawg.nist.gov

August, 21, 2013

AGENDA

- Review NBD-PWG Charter
- Review Subgroups Charter and Deliverables
- Review Overall Workplan
- Subgroup Activities Report
  - Definitions & Taxonomies Subgroup
  - Requirements Subgroup
  - Security & Privacy Subgroup
  - Reference Architecture Subgroup
  - Technology Roadmap Subgroup
- Expected Deliverables – working drafts outline
- Face-to-face Meeting – September 30, at NIST
- Registration: https://www-s.nist.gov/CRS/conf_disclosure.cfm?&conf_id=6552
  - Deliverables Presentation & Discussion
  - Breakout Sessions by Subgroups
  - Announcement for Next Steps
- Q/A

SCOPE (M0001)

The focus of the (NBD-PWG) is to form a community of interest from industry, academia, and government, with the goal of developing a consensus definitions, taxonomies, secure reference architectures, and technology roadmap. The aim is to create vendor-neutral, technology and infrastructure agnostic deliverables to enable big data stakeholders to pick-and-choose best analytics tools for their processing and visualization requirements on the most suitable computing platforms and clusters while allowing value-added from big data service providers and flow of data between the stakeholders in a cohesive and secure manner.

NBD-PWG

SCOPE AND DELIVERABLES

DELIVERABLES: *Working Draft for*

1. *Big Data Definitions*
2. *Big Data Taxonomies*
3. *Big Data Requirements*
4. *Big Data Security and Privacy Requirements*
5. *Big Data Architectures Survey*
6. *Big Data Reference Architectures*
7. *Big Data Security and Privacy Reference Architectures*
8. *Big Data Technology Roadmap*

LAUNCHED DATE:
June 26, 2013

TARGET DATE:
September 27, 2013

SUBGROUPS

AND THEIR SCOPES AND DELIVERABLES

Definitions and Taxonomies
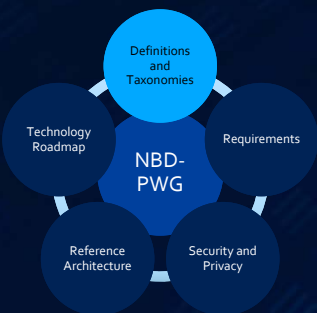
Requirements and Use Cases

Technology Roadmap

NBD-PWG

Reference Architecture

Security and Privacy

Definitions and Taxonomies

Nance Grady, SAIC
Natasha Balac, SDSC
Eugene Luster, R2AD

Scope (M0018)

The focus is to gain a better understanding of the principles of Big Data. It is important to develop a consensus-based common language and vocabulary terms used in Big Data across stakeholders from industry, academia, and government. In addition, it is also critical to identify essential actors with roles and responsibility, and subdivide them into components and sub-components on how they interact/ relate with each other according to their similarities and differences.

Tasks

- For Definitions: Compile terms used from all stakeholders regarding the meaning of Big Data from various standard bodies, domain applications, and diversified operational environments.

- For Taxonomies: Identify key actors with their roles and responsibilities from all stakeholders, categorize them into components and subcomponents based on their similarities and differences

- Develop Big Data Definitions and taxonomies documents

Requirements and Use Cases

Geoffrey Fox, U. Indiana
Joe Paiva, VA
Tsegereda Beyene, Cisco
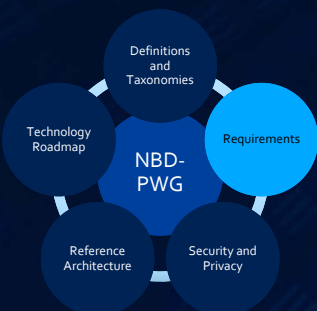
Scope (M0020)

The focus is to form a community of interest from industry, academia, and government, with the goal of developing a consensus list of Big Data requirements across all stakeholders. This includes gathering and understanding various use cases from diversified application domains.

Tasks

- Gather input from all stakeholders regarding Big Data requirements.

- Analyze/prioritize a list of challenging general requirements that may delay or prevent adoption of Big Data deployment

- Develop a comprehensive list of Big Data requirements

Definitions and Taxonomies

Technology Roadmap

NBD-PWG

Requirements

Reference Architecture

Security and Privacy

SUBGROUP

Security and Privacy

Arnab Roy, CSA/Fujitsu
Nancy Landreville, U. MD
Akhil Manchanda, GE
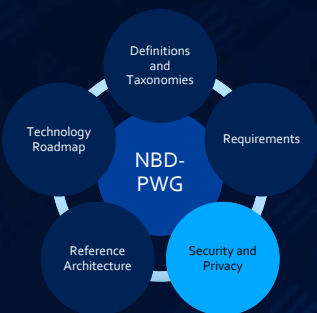
Scope (M0019)

The focus is to form a community of interest from industry, academia, and government, with the goal of developing a consensus secure reference architecture to handle security and privacy issues across all stakeholders. This includes gaining an understanding of what standards are available or under development, as well as identifies which key organizations are working on these standards.

Tasks

- Gather input from all stakeholders regarding security and privacy concerns in Big Data processing, storage, and services.

- Analyze/prioritize a list of challenging security and privacy requirements that may delay or prevent adoption of Big Data deployment

- Develop a Security and Privacy Reference Architecture that supplements the general Big Data Reference Architecture

Reference Architecture

Orit Levin, Microsoft
James Ketner, AT&T
Don Krapohl, Augmented Intelligence

Scope (M0021)

The focus is to form a community of interest from industry, academia, and government, with the goal of developing a consensus-based approach to orchestrate vendor-neutral, technology and infrastructure agnostic for analytics tools and computing environments. The goal is to enable Big Data stakeholders to pick-and-choose technology-agnostic analytics tools for processing and visualization in any computing platform and cluster while allowing value-added from Big Data service providers and the flow of the data between the stakeholders in a cohesive and secure manner.

Tasks

- Gather and study available Big Data architectures representing various stakeholders, different data types,' use cases, and document the architectures using the Big Data taxonomies model based upon the identified actors with their roles and responsibilities.

- Ensure that the developed Big Data reference architecture and the Security and Privacy Reference Architecture correspond and complement each other.

Technology Roadmap

Carl Buffington, USDA/Vistronix
Dan McClary, Oracle
David Boyd, Data Tactic
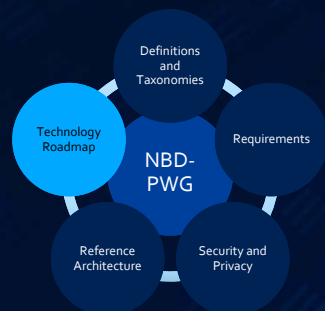
Scope (M0022)

The focus is to form a community of interest from industry, academia, and government, with the goal of developing a consensus vision with recommendations on how Big Data should move forward by performing a good gap analysis through the materials gathered from all other NBD subgroups. This includes setting standardization and adoption priorities through an understanding of what standards are available or under development as part of the recommendations.

Tasks

- Gather input from NBD subgroups and study the taxonomies for the actors' roles and responsibility, use cases and requirements, and secure reference architecture.

- Gain understanding of what standards are available or under development for Big Data

- Perform a thorough gap analysis and document the findings

- Identify what possible barriers may delay or prevent adoption of Big Data

- Document vision and recommendations

NIST Big Data Public Working Group and Subgroups Work Plan

Week	Def. & Tax.	Requirements	Sec. & Privacy	Ref. Arch	Tech. Roadmap
June 26	NBD-PWG (13:00PM – 15:00PM) Kick-off Meeting				
July 3	NBD-PWG (13:00PM – 15:00PM) Establish Subgroups with Co-Chairs, Subgroups Charter, Overall OWG direction				
July 8 - 12	Mondays 10:00AM – 12:00PM	Tuesdays 10:00AM – 12:00PM	Wednesdays 10:00AM – 12:00PM	Thursdays 10:00AM – 12:00PM	Fridays 10:00AM – 12:00PM
July 15 – 19	Definitions & Characteristics	Collect general use cases, identify requirements	Collect security and privacy use cases,	Analyze use cases from Reqs. & Sec. subgroups	Vision Characteristics & Def.
July 22 – 26	Tax.: Roles, activities, components & subcomp.	Categorize reqs., Identify missing reqs.	Identify requirements	Create conceptual model, identify actors,	Taxonomies Roles & Activities
July 24	NBD-WG (13:00PM – 15:00PM) Subgroups report: Sharing and brainstorming results				
July 29 – Aug. 2				Identify usage scenarios, iden. Implement. Scenarios	Use cases & scenarios Ref. Architecture
Aug 5 – 9				Create ref. architecture	Standards & Activities Gap Analysis
Aug 12 – 16					Standardization Priorities ??? Strategy of Adoption
Aug 19 – 23					Strategy of Implement. Resourcing
Aug. 21	NBD-WG (13:00PM – 15:00PM) Subgroups report: Present and Discuss Working Draft Outline				
Aug. 26 - 30					Recommendations
Sept. 2 – 6					
Sept. 4	NBD-WG (13:00PM – 15:00PM) Subgroups report: Present and Discuss Rough Draft				
Sep 9 – 13					
Sep 16 - 20					
Sep 23 – 27					
Sep 25	NBD-WG (13:00PM – 15:00PM) Subgroups report: Present and Discuss Final Draft				
Sep 30	Big Data Workshop, NIST - Deliverables Presentation & Discussion - Breakout Sessions by Subgroups - Announcement for Next Steps				

Events Sequence & Information Flow Between Subgroups

Definition & Taxonomy ②

Security & Privacy ② ①

Reference Architecture ②

Technology Roadmap ③

Requirements & Use cases ①
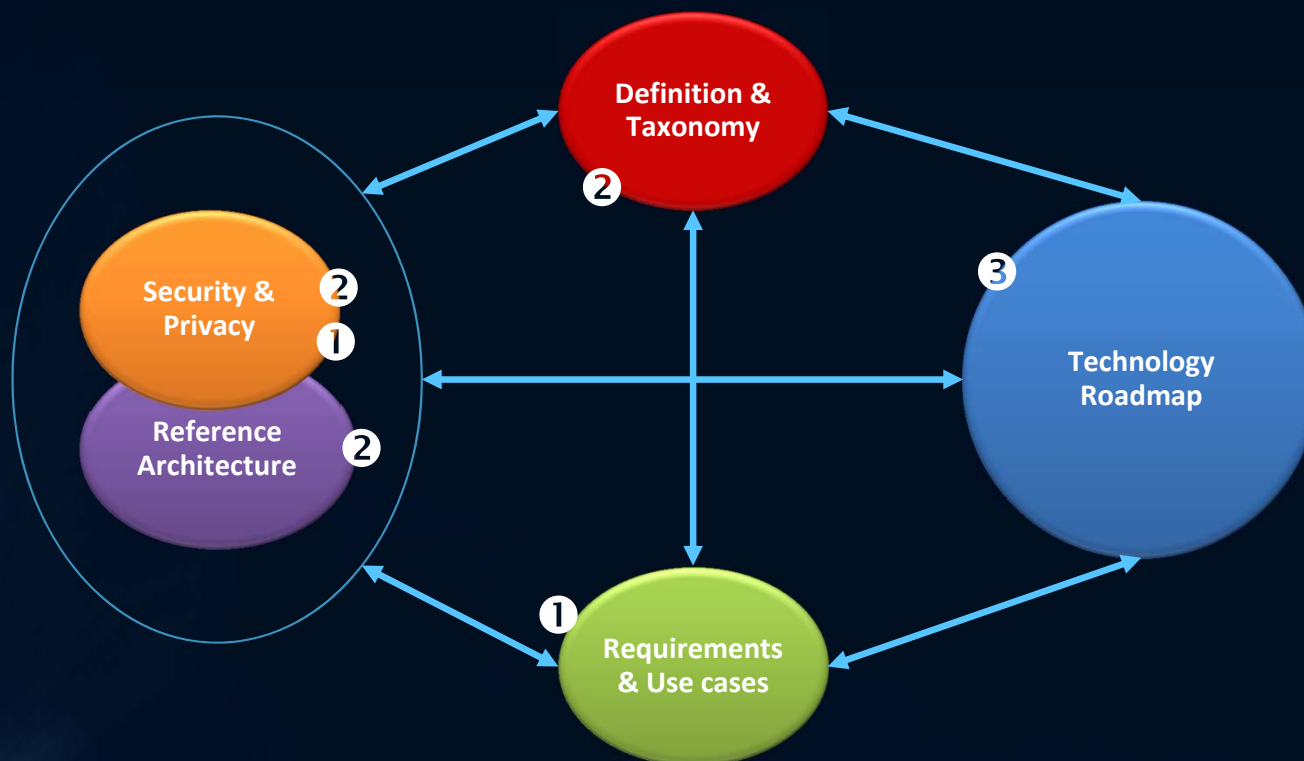
Due to time constraints, activities carry out in parallel.

	July (brainstorm)	Aug (outline)	Sep (write-up)
Definitions and Taxonomies	*(1) For Definitions: Compile terms used from all stakeholders regarding the meaning of Big Data from various environments. (2) For Taxonomies: Identify key actors with their roles and responsibilities from all stakeholders, categorize them into components and subcomponents based on their similarities and differences*		
Requirements and Use Cases	*(1) Gather input from all stakeholders regarding Big Data requirements. (2) Analyze/prioritize a list of challenging general requirements that may delay or prevent adoption of Big Data deployment.* **(16? Use cases received; ~26 general requirements in 7 categories)**		
Security and Privacy	*(1) Gather input from all stakeholders regarding security and privacy concerns in Big Data processing, storage, and services. (2) Analyze/prioritize a list of challenging security and privacy requirements that may delay or prevent adoption of Big Data deployment.*		
Reference Architecture	*(1) Gather and study available Big Data architectures representing various stakeholders, different data types,' use cases, and document the architectures using the Big Data taxonomies model based upon the identified actors with their roles and responsibilities. (2) Ensure that the developed Big Data reference architecture and the Security and Privacy Reference Architecture correspond and complement each other.*		
Technology Roadmap	*(1) Gather input from other subgroups and study the taxonomies for the actors' roles and responsibility, use cases and requirements, and secure reference architecture. (2) Gain understanding of what standards are available or under development for Big Data. (3) Perform a thorough gap analysis and document the findings. (4) Identify what possible barriers may delay or prevent adoption of Big Data.*		

Subgroups Activities Report

- Definitions & Taxonomies (M0024 – ongoing discussion; M0142 – Definitions)
- Requirements and Use Cases (M0105 – use cases; M0125 – requirements; M0135 – working draft)
- Security & Privacy (https://docs.google.com/document/d/1oahT1sTwb7DoCeYoBGwMQy7aUFoiFAR6_rIqMif9m9A/edit?usp=sharing)
- Reference Architecture (M0100 – discussion; M0126 – combined RAs)
- Technology Roadmap
  - M0132 – BD Decision Framework
  - M0122 – BD Technology Readiness
  - M0094 – Initial graphical vision of the Tech Roadmap
  - M0052 – BD Categories
  - Others...

Subgroups Working Draft Outline

- Definitions & Taxonomies (M0142)
- Requirements and Use Cases (M0152)
- Security & Privacy (M0110 – requirements; Moxxx – architecture)
- Reference Architecture (M0151 – white paper; M0123 – working draft)
- Technology Roadmap (M0087)

BIG DATA WORKSHOP (FACE-TO-FACE MEETING):

Date: September 30, 2013
Location: NIST
Registration: https://www-s.nist.gov/CRS/conf_disclosure.cfm?&conf_id=6552