**NIST Big Data Public Working Group (NBD-PWG)**

**NBD-PWD-2016/M0536**

**Source:    NBD-PWG**

**Status:     Draft**

**Title:       NBD-PWG Meeting Agenda for Jun 21, 2016**

**Author:   NBD-PWG Subgroup Co-Chairs**

Meeting logistics

Date/time: June 21, 1:00PM – 3:00PM EDT

Web conferencing tool: <https://global.gotomeeting.com/join/790820565>

Audio: Use your microphone and speakers (VoIP) - a headset is recommended, or, call in using your telephone (US, long distance): +1 (646) 749-3122, access code/meeting ID: 790-820-565

## **Agenda**

1. Tentative face-to-face NBD-PWG workshop at NIST, Dec. 14 – 16, 2016 (tentative)
2. Tentative timelines for our V2 are (A BIG assumption that we have sufficient V2 content):

July 18: public announcement for the Dec. 14-16 workshop on Big Data Interop. Framework

Nov. 14: make available rough DRAFT-1 for Dec. workshop discussion

Dec. 14 – 16: NBDIF workshop (could be two days or two and half days)

Dec. 19: incorporate all input and produce DRAFT-2 for public comments

Feb. 28, 2017: begin DRAFT-2 for public comment for 45 days; comment close by ~mid-April

Mid-April: incorporate all comments and produce DRAFT-3 for ITL review

Mid-May: incorporate ITL comments and produce DRAFT-4 for WERB (NIST editor board) review

Mid-June: begin WERM review

Mid-Aug: incorporate WERB comments and produce V2 release

1. Continue Subgroup Co-Chairs report – Draft V2 outline/content:

(Our V2 drafts are on Google Drive: Anyone with links can review and download. For online editing, you need an GMAIL account. If interested, email Wo for “edit” access)

1. **Vol. 1 Definitions, Nancy**

Draft:

<https://drive.google.com/open?id=1AFFxut4nu1Yj8uxnzUQoBWW0QwKbcmtqVxnNmU2v0EM>

1. Enhance existing V1 definitions
2. Defining the different patterns of communications between Big Data resources to better clarify the different approaches being taken;
3. Improving the discussions of governance, value, and data ownership; (Tim, Quyen)
4. Developing the Management section; (Scope? Obviously systems, etc.)
5. Developing the Security and Privacy section; (Arnab, Mark)
6. Improve discussion of new behavior (new design patterns?) in big data
   1. Bob IoT
   2. ???
7. Whitepaper: Implication of Big Data in social, business, public, knowledge
8. Architectures (briefly introduce, use GFox descriptions, etc. Brief here, fleshed out elsewhere in Taxonomy.)
9. Orchestrator / Orchestration: Refer to additional detail in the MindMap, (a discussion ensued re: Orchestrator as role, but also as conceptual underpinning for enterprise (“business”) process in the RA.
10. Cross references to other volumes…
11. Discussion of relationship to cloud
12. Discussion of Machine Learning?
13. Do we need a separate section for analytics, or should that be rolled up into the Data Science section?
14. Rewrite “paradigm” section to convert to Frank’s Big Data Field, or “when do you have a big data problem?” (Nancy)
15. **Vol. 2 Taxonomies, Nancy**

Mindmap: <https://www.mindmeister.com/322462463> [Email Wo for access]

Draft: [coming soon]

1. Align with the other v1 documents
2. The Subgroup is continuing to explore the changes in both Management and in Security and Privacy. As changes in the activities within these roles are clarified, the taxonomy will be developed further.
3. In addition, a fuller understanding of Big Data and its technologies should consider the interactions between the characteristics of the data and the desired methods in both technique and time window for performance. These characteristics drive the application and the choice of tools to meet system requirements. Investigation of the interfaces between data characteristics and technologies is a continuing task for the NBD-PWG Definitions and Taxonomy Subgroup and the NBD-PWG Reference Architecture Subgroup.
4. Finally, societal impact issues have not yet been fully explored. There are a number of overarching issues in the implications of Big Data, such as data ownership and data governance, which need more examination. Big Data is a rapidly evolving field, and the initial discussion presented in this volume must be considered a work in progress.
5. Explore the taxonomy’s ability to work with Geoffrey’s blend of HPC and Big Data
6. Cloud Taxonomy: <http://collaborate.nist.gov/twiki-cloud-computing/bin/view/CloudComputing/ReferenceArchitectureTaxonomy?sortcol=5;table=4;up=1#sorted_table>
7. Explore formal methods to define taxonomy
8. Efforts should leverage definitions produced elsewhere to avoid ocean-boiling or unresolvable concerns that are less salient for big data.
9. **Vol. 3 Use Cases & Requirements, Geoffrey and Piyush**

UC2 Google Form: <http://bit.ly/1ff7iM9>

(Mark is willing to take time to interview people who have BD applications)

Students’ Project: <https://docs.google.com/document/d/1OCPO2uqOkADvoxynRyZwh5IyFQ2_m1fkpBVMo3UBblg/>

Draft: [coming soon]

1. Review, finalize, and begin collecting new use cases based from UC Template V2
2. How to analyze new use cases with SnP info? How to coordinate vol. 4?
3. Draw on the use case classification to suggest classes of software models and system architectures
4. A more detailed analysis of reference architecture based on sample codes that are being implemented in a university class.
5. Collect benchmarks that capture the “essence” of individual use cases.
6. Additional work may arise from these or other NBD-PWG activities. Other future work may include collection and classification of additional use cases in areas that would benefit from additional entries, such as Government Operations, Commercial, Internet of Things, and Energy. Additional information on current or new use cases may become available, including associated figures. In future use cases, more quantitative specifications could be made, including more precise and uniform recording of data volume. In addition, further requirements analysis can be performed now that the reference architecture is more mature.
7. **Vol. 4 Security & Privacy, Arnab and Mark**

Topics Assignment:

<https://drive.google.com/open?id=1Rs6LkwMkNv-116tGFRg0OJU1pi2FZXsMlFMRBpBtZ9g>

Unified SnP Taxonomy:

Visio format: <https://drive.google.com/open?id=0B6JGdUJOefinNldMWEZZX1J1YWs>

PDF format: <https://drive.google.com/open?id=0B6JGdUJOefincjNQTFBfZEQyVUU>

1. How to coordinate with vol.3 new use cases on SnP?
2. Developing the unified security and privacy taxonomy:
3. Exploring governance, risk management, data ownership, and valuation with respect to Big Data ecosystem, with a focus on security and privacy;
4. Contextualizing the content of Appendix B in the NBDRA; and
5. Expanding the privacy discussion within the scope of this volume;
6. Exploring privacy in actionable terms based on frameworks such as those described in NISTIR 8062 [7] with respect to the NBDRA.
7. Set of whitepapers
8. Vol. 5 White Paper Survey, No Future Work for now
9. **Vol. 6 Reference Architecture, David**

Draft: <https://drive.google.com/open?id=1Tol4tKxo-xugIjVyjfm9GSJpbNRFB3AFXQ4unKJjW0w>

Activities/Functional Diagrams: <https://drive.google.com/open?id=0B4Y1OQqW13RLcG90ZXN3dHJiWjg>

White Paper Sample:

<https://drive.google.com/open?id=1_4o2vN8kNInc2e08aySgRBqU-dMizShbaz8_DJ6wiWM>

1. Reference Architecture Refinement
2. Establish activity and functional component views beyond the current conceptual view
3. Define high level and general activities and functional components within each view
4. Identify high level stakeholder and map their concerns to activities and functional components
5. Reference Architecture application
6. Establish white paper template
7. Implement the NIST identified six use cases and/or other use cases from the 62 (51 generals and 11 security and privacy) collected use cases or others
8. Identify development environment for hosting the use case implementations
9. Create white papers by working with domain experts to identify workflow and interactions among the NBDRA components and fabrics
10. Review, analyze white papers high-level interactions and workflows and aggregate them into preliminary general interfaces
11. Conformance approach
12. **Vol. 7 Standards Roadmap, Russell**

Draft: [coming soon]

1. Examine all version 1 volumes and:
2. Identify available standards, and those under development.
3. Gap the differences between version 1 volumes and list of standards, and continue to build and refine the gap analysis and document the findings;
4. Extend related SDO listing and establish criteria on how to select relevant SDOs
5. Support harmonizing of terminology between volumes.
6. Identify early adopters from academia, government and industry.
7. Identify barriers to big data adoption.
8. Identify where standards may accelerate the adoption and interoperability of Big Data technologies;
9. Further map standards to NBDRA components and the interfaces between them.
10. Enhance gap analysis on how to enable the RA
11. Document vision and recommendations for future standards activities
12. Engage communities to attract additional NBD-PWG participation and drive consensus on how big data should move forward.