**Contributions for NBDIF Version 3 Topics**

Version 3 work on the National Institute of Standards and Technology (NIST) Big Data Interoperability Framework (NBDIF) will focus on the enhancement of the NIST Big Data Reference Architecture (NBDRA) interfaces, validation of the interfaces, and improvement of several topics identified during version 2 work (in Section 1.5 of each version 2 volume). The subgroup cochairs and the NIST Big Data Public Working Group (NBDPWG) members indicated that the topics may not be as fully developed as they desired. The development of these topics during version 3 depends entirely on the submission of text from the NBDPWG and public participants.

To submit text for a particular topic, please follow the steps below:

1. Select a topic from a list below.
2. Write text. Include complete references to support the text. Discuss specifics of the topic with the subgroup cochair(s).
3. Submit the text using the “NBDIF Version 3 Topic Submission Template” to Wo Chang at wchang@nist.gov. If the text is submitted by Friday 9AM Eastern Time, the submission will be on the agenda for the following Tuesday call (depending on time availability in the agenda).
4. Wo will post the text submission to the NBDPWG portal (<https://bigdatawg.nist.gov>) by Friday 5PM Eastern Time.
5. During the Tuesday call, the NBDPWG will discuss the text. Please plan to attend the Tuesday meeting when your text is scheduled. If you have any scheduling conflicts or questions, please contact Wo Chang.

To submit a new topic that is not on the list please do the following:

* Submit the new topic to Wo by Friday for inclusion on the following Tuesday meeting agenda. The new topic will be discussed by the NBD-PWG to decide if it is within scope.
* Within 60 days from the acceptance date of the topic, submit the text to include in the identified volume following the steps outlined above.

Important dates:

* March 30: Deadline to submit new topics to be considered for version 3
* June 1: Deadline for submission of any text to be included in version 3
* August 10: Announcement for 45 days public comment period
* September 23: Public comment period ends

Volume 1

| Topic # | Volume 1 Topic Description | Possible related section(s) |
| --- | --- | --- |
|  | Better characterization of implementation differences in instances of NoSQL (Not Only or No Structured Query Language [SQL]) or NewSQL data platforms | 4.2.2, 4.2.3, |
|  | Guidance on system and data flow metrics | 5.6 |
|  | Discussion of changes in algorithms and statistical measures | 5.1, 5.6 |
|  | Discussions of new programming languages | 4.3.5 |
|  | Expanding relationships to machine learning and virtual reality | 4.3 |
|  | Add new citations |  |

Volume 2

|  |  |  |
| --- | --- | --- |
| Topic # | Volume 2 Topic Description | Possible related section(s) |
| 2.1 | NIST cloud taxonomy | New section |
| 2.2 | Security and privacy taxonomy with respect to NIST SP800-53 |  |
| 2.3 | Other relevant taxonomies | New section |

Volume 3

|  |  |  |
| --- | --- | --- |
| Topic # | Volume 3 Topic Description | Possible related section(s) |
| 3.1 | Submit new use cases using the Use Case Template 2 (https://bigdatawg.nist.gov/\_uploadfiles/M0621\_v2\_7345181325.pdf) |  |
| 3.2 | Evaluate additional use cases as they are submitted |  |

Volume 4

| Topic # | Volume 4 Topic Description | Possible related section(s) |
| --- | --- | --- |
| 4.1 | The NBD-PWG has expanded Big Data security and privacy analysis and fostered adoption through deeper cross-linking to related standards. Progress was made but work remains. |  |
| 4.2 | Explicit phase-specific guidance was initiated, but not fully implemented, especially in Appendix A. |  |
| 4.3 | Develop improved guidelines for integrating supporting Big Data systems dedicated to security and privacy (Big Data security and privacy dogfooding).[[1]](#footnote-1) Healthcare is the strongest use case, where risks are best understood. |  |
| 4.4 | Incorporate security and privacy metadata-rich Big Data orchestration processes, enabled by tools such as Rundeck (Spinellis, 2014). |  |
| 4.5 | Facilitate incorporation of security and privacy models for the software development life cycle. For example, there is a need to describe Big Data security and privacy practices that address the full life cycle of Big Data analytics and machine learning. |  |
| 4.6 | Draft a Big Data-annotated version of the NIST Privacy Catalog (see (National Institute of Standards and Technology (NIST), 2014), Appendix J). |  |
| 4.7 | Identify Big Data touchpoints for Privacy by Design, Organisation for Economic Co-Operation and Development (OECD), and other external privacy guidelines. |  |
| 4.8 | Integrate models such as Sensing as a Service (Zaslavsky, Perera, & Georgakopoulos, 2012) |  |
| 4.9 | Provide a deeper explanation of Big Data Application Provider security and privacy requirements. |  |
| 4.10 | Develop security and privacy risk frameworks for specific design patterns (apart from cloud), including distributed computing, middleware (enterprise service bus), agent-based, recommendation engines, and web portals fronting legacy applications. |  |
| 4.11 | More clearly identify where Big Data systems management intersects with security and privacy guidelines. The gold standard use case is the use of logging data for both operational intelligence and security and privacy, though the mapping is demonstrably nonorthogonal. |  |
| 4.12 | Depict security and privacy policy and metadata orchestration using descriptions of test beds, such as those developed in the *NBDIF: Volume 8, Reference Architecture Interfaces* document. |  |
| 4.13 | Update or build new frameworks for Big Data that reference existing International Organization for Standardization (ISO) and other standards for Big Data life cycle, audit, configuration management, and privacy preserving practices. |  |
| 4.14 | Contextualize the content of Appendix C (Internal Security Considerations within Cloud Ecosystems) in the NBDRA. |  |
| 4.15 | Create additional mapping for the use cases to the NBDRA and security and privacy taxonomy. |  |
| 4.16 | Enhance the discussion of infrastructure management, including in-depth exploration of *left shift* and its implications for Big Data security and privacy, implications for infrastructure as code, and relevance to NIST Critical Infrastructure. |  |
| 4.17 | Identify Big Data challenges associated with ownership traceability, custody and curation – especially in light of data volatility introduced by mergers, cross-border flows and project life cycle (Section 2.3.5). |  |
| 4.18 | Enhance Effects of Emerging Technology on Big Data Security and Privacy | 2.4 |
| 4.19 | Enhance Cloud Computing | 2.4.1 |
| 4.20 | Enhance Internet of Things and CPS | 2.4.2 |
| 4.21 | Enhance Big Data Security Safety Annotation | 2.4.7.3 |
| 4.22 | Enhance Orchestration in Weak Federation Scenarios | 2.4.7.5 |
| 4.23 | Enhance Consent and the Glass-breaking Scenario | 2.4.7.6 |
| 4.24 | Enhance Security and Privacy Approaches in Analytics | 5.4 |
| 4.25 | Enhance Risk Management | 5.6 |
| 4.26 | Enhance Big Data Security Modeling and Simulation (ModSim) | 5.7 |
| 4.27 | Enhance Packet-Level Traceability / Reproducibility | 7.1 |
| 4.28 | Develop SABSA and Zachman Framework | Could be new subsection in section 8 |
| 4.29 | Enhance Configuration Management for Big Data | 8.3 |
| 4.30 | Enhance Dependency Models | 8.3.2 |
| 4.31 | Enhance Big Data SDLC Standards and Guidelines | 8.4 |
| 4.32 | Enhance Security and Privacy Events in Application Release Management | 8.4.1.2 |
| 4.33 | Enhance Orchestration | 8.4.1.3 |
| 4.34 | Enhance ‘Least Exposure’ Big Data Practices | 8.4.6.2 |
| 4.35 | Enhance Logging | 8.4.6.3 |
| 4.36 | Enhance Big Data Governance | 8.5 |
| 4.37 | Enhance Emerging Technologies | 8.6 |
| 4.38 | Enhance Machine Learning, AI, and Analytics for Big Data Security and Privacy | 8.6.2 |

Volume 6

| Topic # | Volume 6 Topic Description | Possible related section(s) |
| --- | --- | --- |
| 6.1 | **Develop additional architecture views** and associated descriptions to reflect additional areas of concern | 4 |
| 6.2 | **Develop architecture views focused on addressing data concerns.** The current views described in this document focus heavily on the process and software components of the architecture. Since, this reference architecture deals with Big Data, additional views focused on addressing data concerns should be developed and incorporated into this architecture | 4 |
| 6.3 | **Develop data centric sub-views** focused on those Big Data characteristics (e.g., volume, velocity, variety)would likely be appropriate | 4 |
| 6.4 | **Develop data and interface oriented views.** In addition, the highly connected nature of today’s systems and the challenges associated with moving data with Big Data characteristics between systems tend to drive system-of-systems approaches as solutions to many problem areas. This means, that architectures describing such solutions will need to be able to effectively address the interface requirements between Big Data systems and their components, thus requiring data and interface oriented views. | 4 |
| 6.5 | **Develop security and management oriented views.** Finally, the critical nature of securing and managing such architectures reflected in the security and management fabrics of the reference architecture will require views specifically addressing those concerns so that the solutions can be effectively documented and validated. | 4 |

Volume 7

| Topic # | Volume 7 Topic Description | Possible related section(s) |
| --- | --- | --- |
| 7.1 | Document recommendations for future standards activities | 4 |
| 7.2 | Further map standards to NBDRA components and the interfaces between them | Appendix C |
| 7.3 | Map additional requirements to standards | 3.1.1 |
| 7.4 | Map additional use cases to standards (e.g., use cases 2, 6, 34) | 3.1.2 |
| 7.5 | Exploration of the divergence of technologies and common project methodologies and the impact on standards creation |  |
| 7.6 | Investigate the impact of standards for IoT, including a recognized need in the area of encrypted network traffic | New section |
| 7.7 | Consider the need for standards in the areas of network connectivity, complex event processing, PaaS, and crowdsourced mediation | New section |
| 7.8 | Explore existing gaps in data standards, including topics such as types of data sets, application level services, open data, and government initiatives | 4.1, 4.2 |
| 7.9 | Commercial datasets and open marketplaces |  |
| 7.10 | Construct gap closure strategies | 4.2 |

Volume 8

| Topic # | Volume 8 Topic Description | Possible related section(s) |
| --- | --- | --- |
| 8.1 | Continue development of reference software | All |

Volume 9

| Topic # | Volume 9 Topic Description | Possible related section(s) |
| --- | --- | --- |
| 9.1 | Areas where standards could have significant impact |  |
| 9.2 | Technical challenges with data integration and preparation, specifically dealing with variables of different magnitudes | 3.2, 4.4 |
| 9.3 | Pathways for organizations to modernize to facilitate the successful transition from existing systems to more modern systems | 4.1, 4.2 |

1. Typically such supporting security and privacy Big Data is provided as part of a fully integrated Build Phase, but some solutions can implement Security as a Service, with some or all security and privacy resources provided by third parties. Third parties may specialize in security and privacy for specific domains, with machine learning, ontologies, and other specialized resources that may be beyond the capabilities of Build architects. [↑](#footnote-ref-1)