**NIST BD joint subgroup meeting minutes for Nov. 25, 2013**

Prepared by Sanjay Mishra, Verizon

**Agenda**

* Recap Current Deliverables and Share Future Activities

Current Deliverables:

* + V1.0 Documents for Requirements and Sec. & Privacy – for technical editing
  + V1.0 Documents for Def. & Tax., Ref. Arch., and Tech. Roadmap – need harmonization
  + Schedule: every Monday, 1:00PM – 4:00PM EDT till Dec. 23, 2013
  + Goal: finalize all V1.0 draft documents (ready for technical editing) **by Dec. 23, 2013**
  + Reflector: please use [bigdatadeftax@nist.gov](mailto:bigdatadeftax@nist.gov), [bigdataarch@nist.gov](mailto:bigdataarch@nist.gov), and/or [bigdataroadmap@nist.gov](mailto:bigdataroadmap@nist.gov) for discussion

**Definition and Taxonomy**

* **Definitions**
  + Nancy reported section 2.1 on Big Data Paradigm, Big Data and Big Data Engineering have been updated.
  + Comments over the call:
    - (Phil) How is computational Complexity handled in Big data? (Traveling Salesman problem). Issue with Big Data is more you dig, more data you get…. Focus should be on what problems can be solved and what cannot be solved by Big Data.
    - (Wo) Do we need separate sections on Big Data Engineering and Data Science & Data Scientist? Also, should Big Data Engineering cover the storages and networking distributed data source sites?
    - (Bernard) For Architecture, have not captured terms such as splitters, combiners and Hadoop (characteristics requirements are same even if you don’t use “Hadoop”).
* **Taxonomy**
  + The group went over the taxonomy both from the document and also the online mind map. The strategy is to identify 1st level of activities from the actors so that the aggregated general activities along with the actors can be transferred to the Reference Architecture write-up. From the write-up point of view, it would make better sense to combine Taxonomy into the Reference Architecture document.

**Reference Architecture**

Changes to the document

1. Orit reported specific updates to section 2, 3 and 4 have been made (need to enable track changes to easily identify changes)
2. Section 2 last paragraph: if this paragraph expands with more substantial text, it would be good to have its own section and called it Abstraction Services. The idea is to demonstrate to reader how the RA key components interact with each other.
3. Section 3: The overall structure is good which includes actors and activities but it needs to harmonize with the taxonomy part. It will be good to include text from the Taxonomy to make it more complete and harmonize.
4. Section 4: Thanks to Felix to start the Section 4.1 on Management. Felix will expands that section to include other management functionalities since managing big data computing cluster is different from traditional client/server practices.

**Technology Roadmap**

* Pw Carey has provided minor input to Section 5 of TR document. The changes were not reviewed on call as Pw had dropped at 3PM. Pw is still reviewing Section 7 and will make comments before next meeting.
* David will ping the TR authors for sections that are independent from subgroups and make report for next meeting.

**General**

* Anyone interested in participating to the ISO/IEC JTC1 Big Data Study Group (BD-SG), please send an e-mail to Wo.
* The 1st ISO/IEC JTC1 BD-SG meeting in the US is tentatively scheduled for Jan’28 to Jan’31, 2014.

**Action Items**

1. Nancy will take Taxonomy text and merge into section 3 of Reference architecture
2. System Management needs (section 4.1) updated (Owner: Felix Njeh)
3. Need a synch up with Security (Owner: Orit)
4. All participants should review mind-map to see what is missing, from actors, and activities perspective. (Owner(s): All) For accessing mind-map, please send email to Wo.
5. Pw will review and comment TR Section 7 before next meeting.
6. David will progress report for the sections not depending from other subgroups.

**Online Attendee List**

1. Wo Change
2. Nancy Grady
3. Rick Jones
4. William Miller
5. PW Carey
6. Keith Hare
7. Quyen Nguyen
8. Orit Levin
9. Phil Journeau
10. Panagiotis Papadopoulos
11. Bev Corwin
12. Yuri Demchenko
13. Pavithra Kenjige
14. David Boyd
15. Felxi Njeh
16. Ian Gorton
17. Sanjay Mishra

**Online Chat log**

(1:04 PM) Pw Carey, Compliance Partners, LLC joined.

(1:04 PM) Orit joined.

(1:05 PM) Rick Jones (Joseki Group LLC) disconnected.

(1:05 PM) Rick Jones (Joseki Group LLC) joined.

(1:05 PM) Pw Carey, Compliance Partners, LLC: Ok...we can hear you quite well....Pw

(1:07 PM) Bev Corwin joined.

(1:07 PM) Phil joined.

(1:08 PM) William Miller joined.

(1:11 PM) Pw Carey, Compliance Partners, LLC: Is it possible to turn-on Tracking Changes in a beautiful rainbow of colors.....? just asking...Respectfully yours, Pw

(1:13 PM) Pw Carey, Compliance Partners, LLC: What about changing."The answer is an unsatisfying "it depends".....to "...depending upon the final use and requirements of the data...."

(1:14 PM) Pw Carey, Compliance Partners, LLC: as Big Data is good for some things but not for all.....

(1:17 PM) Nancy Grady (SAIC): I would very much appreciate if folks would turn on track changes and send me updates. I can compile and get a new version out

(1:18 PM) Pw Carey, Compliance Partners, LLC: Dear Nancy...going forward we will turn on Track Changes each time we comment on a document....thanks for reminding us....Respectfully yours, Pw

(1:19 PM) Pw Carey, Compliance Partners, LLC: Correct, Big Data should not be considered a 'Secret Sauce' for data handling.....Pw

(1:20 PM) Phil disconnected.

(1:21 PM) Pw Carey, Compliance Partners, LLC: Good point.....

(1:22 PM) Dr. Panagiotis Papadopoulos EDF Energy R&D UK Centre joined.

(1:22 PM) Pw Carey, Compliance Partners, LLC: How do you determine when the best solution is Big Data....?

(1:23 PM) Pw Carey, Compliance Partners, LLC: When the data is 'regulated' and folks can and do go to jail for failure to ensure the accuracy of the data in question.....perhaps....no?

(1:24 PM) Pw Carey, Compliance Partners, LLC: Could be one criteria...for choosing a Big Data solution...?

(1:25 PM) William Miller disconnected.

(1:25 PM) Yuri Demchenko 9UvA) joined.

(1:26 PM) David Boyd joined.

(1:27 PM) Pw Carey, Compliance Partners, LLC: We're on mutae and can hear you, too, too....Pw

(1:29 PM) Pw Carey, Compliance Partners, LLC: Data Storage Networking can be a factor...data leakage....data borders...et al....

(1:29 PM) Keith Hare, JCC Consulting, Inc. joined.

(1:30 PM) Orit disconnected.

(1:32 PM) Pw Carey, Compliance Partners, LLC: We're coming from a position of an Auditor, Forensics Auditor, GRC Auditor...et cetera....for detecting Fraud in all it's forms....Pw

(1:32 PM) Quyen Nguyen joined.

(1:33 PM) Orit Levin (Microsoft) joined.

(1:33 PM) Pw Carey, Compliance Partners, LLC: Treat Big Data as a Single Document.....yes...?

(1:35 PM) Pw Carey, Compliance Partners, LLC: One clean version of a Mind Map...

(1:36 PM) Pw Carey, Compliance Partners, LLC: Consider Taxonomy & Architecture together.....?

(1:37 PM) Pw Carey, Compliance Partners, LLC: Pw also, is working on Section 7, too.....

(1:38 PM) Pw Carey, Compliance Partners, LLC: Splitters, Binders....parts of how the architecture works....

(1:39 PM) Pw Carey, Compliance Partners, LLC: Good point.....

(1:39 PM) Pw Carey, Compliance Partners, LLC: Rather, good points...

(1:39 PM) Rick Jones (Joseki Group LLC): A reference architecture is supposed to be more high-level than that.

(1:40 PM) PavithraKenjige joined.

(1:40 PM) Pw Carey, Compliance Partners, LLC: We should loose your comments...sir...Respectfully yours, Pw

(1:41 PM) Pw Carey, Compliance Partners, LLC: Rather we should NOT loose your comments...sir

(1:41 PM) Pw Carey, Compliance Partners, LLC: The decision to act upon them is another question....Respectfully yours, Pw

(1:42 PM) David Boyd disconnected.

(1:43 PM) David Boyd joined.

(1:43 PM) Rick Jones (Joseki Group LLC): The abstract pattern that Hadoop represents is divide and conquer. That's about the level of discussion that should be mentioned in a reference architecture.

(1:43 PM) David Boyd disconnected.

(1:43 PM) David Boyd joined.

(1:46 PM) Pw Carey, Compliance Partners, LLC: The use and reference of 'patterns' within our descriptions of Big Data.....too....

(1:47 PM) David Boyd disconnected.

(1:48 PM) David Boyd joined.

(1:51 PM) Felix Njeh (COMINT) joined.

(1:54 PM) Bev Corwin disconnected.

(1:54 PM) Bev Corwin joined.

(1:55 PM) Pw Carey, Compliance Partners, LLC: We reviewed & suggested minor comments to Orit's M0226 document, which we'll send along to her after today's meeting...and yes, we will turn-on Track Changes, too....Pw

(1:56 PM) Pw Carey, Compliance Partners, LLC: We don't trust-um.....Pw

(1:57 PM) Pw Carey, Compliance Partners, LLC: When is our next meeting....?

(1:57 PM) Bev Corwin disconnected.

(1:57 PM) Bev Corwin joined.

(1:59 PM) Felix Njeh (COMINT): I can write

(2:00 PM) Pw Carey, Compliance Partners, LLC: Dear Felx: Congratulatons!!!! Respectfully yours, Pw

(2:00 PM) Felix Njeh (COMINT): Sure, I can add more information

(2:00 PM) Ian G, CMU SEI joined.

(2:00 PM) Felix Njeh (COMINT): Yes

(2:00 PM) Pw Carey, Compliance Partners, LLC: That should read: Dear Felix:....

(2:00 PM) Felix Njeh (COMINT): Yes Sir

(2:00 PM) Pw Carey, Compliance Partners, LLC: & Congratulations....

(2:01 PM) Pw Carey, Compliance Partners, LLC: Also, Felix...if you need any help, we'll be happy to assist....too...Respectfully yours, Pw

(2:01 PM) Felix Njeh (COMINT): Is the master document located where I can add the updates?

(2:01 PM) Felix Njeh (COMINT): Thanks PW

(2:02 PM) Orit Levin (Microsoft): The master document is 226v9

(2:02 PM) Pw Carey, Compliance Partners, LLC: You're welcome and where is the Master Document you refer to....Pw

(2:02 PM) Felix Njeh (COMINT): @PW, we can definitely work together on that portion of the doc. Thanks

(2:02 PM) Pw Carey, Compliance Partners, LLC: Thanks Orit....Pw

(2:02 PM) Pw Carey, Compliance Partners, LLC: & Thank you Felix...too...

(2:04 PM) Pw Carey, Compliance Partners, LLC: Should there be Audit Architects, too...?

(2:04 PM) Pw Carey, Compliance Partners, LLC: Included within the Mind Map....?

(2:05 PM) Dr. Panagiotis Papadopoulos EDF Energy R&D UK Centre disconnected.

(2:06 PM) Pw Carey, Compliance Partners, LLC: The Mind Map & The Ref. Architecture have been worked in parallel...understood, thanks Pw

(2:25 PM) Yuri Demchenko 9UvA) disconnected.

(2:25 PM) Bev Corwin disconnected.

(2:26 PM) Pw Carey, Compliance Partners, LLC: Felix my email address is...pwc.pwcarey@gmail.com and we'll follow-up with you later this week.....thanks and have a nice and relaxing Thanksgiving....Respectfully yours, Pw

(2:35 PM) Rick Jones (Joseki Group LLC) disconnected.

(2:45 PM) Orit Levin (Microsoft): Thank you, Nancy!!!

(2:50 PM) Pw Carey, Compliance Partners, LLC: Dear Sir: Could you take a couple of minutes and jot down your thoughts which you brought up today...no more than five minutes time and send them to me, please...Respectfully yours, Pw - pwc.pwcarey@gmail.com....Respectfully yours, Pw

(2:50 PM) Pw Carey, Compliance Partners, LLC: & don't worry about spelling and/or puncutation....too...Pw

(2:52 PM) Pw Carey, Compliance Partners, LLC: Yes, identifying the Key Components....

(2:53 PM) Felix Njeh (COMINT): I see examples of Data Providers - Sensors is missing?

(2:54 PM) Pw Carey, Compliance Partners, LLC: Dear Felix...share any others you flag, please...Pw

(2:56 PM) Pw Carey, Compliance Partners, LLC: A Service Abstraction example would be .... what....?

(2:57 PM) Pw Carey, Compliance Partners, LLC: At a very high level what does the Hypervisor do for the Service Provider...yes?

(2:59 PM) Ian G, CMU SEI: I can't hear Orit at all

(3:00 PM) Pw Carey, Compliance Partners, LLC: I think it's a Microsoft Requirement....perhaps....Pw

(3:00 PM) Pw Carey, Compliance Partners, LLC: No...that can't be...

(3:02 PM) Pw Carey, Compliance Partners, LLC: Dear Fellow Team Members: We must leave but we'll take good virtual notes for the rest of the meeting....In closing, have a nice Thanksgiving....Respectfully yours, Pw

(3:04 PM) Pw Carey, Compliance Partners, LLC disconnected.

(3:06 PM) Ian G, CMU SEI: Wo is describing a role that something like Puppet/Chef would be used for

(3:12 PM) Felix Njeh (COMINT) disconnected.

(3:19 PM) William Miller joined.

(3:19 PM) Keith Hare, JCC Consulting, Inc.: I will participate in the JTC1 Big Data Study group as much as possible.