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

**NBD-PWD-2015/M0439**

**Source: NBD-PWG**

**Status: Draft**

**Title: Web chat and note from Meeting of Sept. 8, 2015**

**Action Items:**

1. Thanks to those willing to apply the NIST identified six use cases

(<http://bigdatawg.nist.gov/_uploadfiles/M0399_v3_9386924556.doc>) to NIST RA Conceptual Model (<http://bigdatawg.nist.gov/_uploadfiles/M0437_v1_8937283419.pptx>) for RA Activities View and RA Functional View. The NIST six use cases are:

* 1. UC#1: Fingerprint Matching
  2. UC#2: Human Face Detection from Video (simulated streaming data)
  3. UC#3: Live Twitter Analysis
  4. UC#4: Big data Analytics for Healthcare Data/Health informatics
  5. UC#5: Spatial Big data/Spatial Statistics/Geographic Information Systems
  6. UC#6: Data Warehouse and Data Mining

The following volunteers are willing to walk through the selected use case with the NIST RA conceptual model:

1. Ann Racuya-Robbins: UC#2
2. Korin Reid: UC#4 and UC#6
3. Russell: UC#3, UC#4, UC#6
4. Gretchen: UC#3
5. PW: UC#3
6. Robert Reyling and Russell will help Dave to work on the Reference Architecture out from IEEE 42010
7. Tim will help Wo to draft some justification to apply free computing in a cloud from AWS, Microsoft, Google, etc.

**Chat Log D:\\_wo\1DMG\2015\\_BigDataWG\Docs\ChatLog 2015\_09\_08 14\_46.rtf**

**Russell (to Everyone)**: 1:27 PM: Russell present and accounted tarde

**Russell (to Everyone)**: 1:28 PM: Trying to get audio on a new machine / no audio yet

**Tim Zimmerlin (to Everyone)**: 1:41 PM: Historically, CORBA is a standard that relates closely to our NBD-PWG architecture. I do not want to use CORBA; it is obsolete.

**Tim Zimmerlin (to Everyone)**: 1:42 PM: CORBA IIORB is a candidate starting point.

**Tim Zimmerlin (to Everyone)**: 1:42 PM: Otherwise, ssh, REST, and SOA are also candidates.

**Tim Zimmerlin (to Everyone)**: 1:43 PM: The current block diagrams are fine as far as a top level POV; however, they leave out a lot of details.

**Tim Zimmerlin (to Everyone)**: 1:46 PM: CloudLab and Chameleon are NSF big data projects with very large server grids. They support user identity, HW reservations, storage reservation with thin provisioning, etc.

**Tim Zimmerlin (to Everyone)**: 1:51 PM: Amazon AWS, Microsoft Azure, and Google Compute Engine are hyperclouds with extensive architectures. E.g., they use custom networks with custom HW devices and NIC cards.

**Tim Zimmerlin (to Everyone)**: 1:53 PM: Dave is talking about "information lifecycle management" in its many incarnations for a broad spectrum of stakeholders and actors. Provenance and curation are critical during the early staging during ingestion and cleansing.

**Tim Zimmerlin (to Everyone)**: 1:56 PM: Suggest: information lifecycle management view, data discovery view ala query/search/filter, pattern discovery view ala machine learning, etc.

**Tim Zimmerlin (to Everyone)**: 1:57 PM: Also suggest: curation and data remediation view.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:58 PM: Pw Carey has Arrived....Yea!!!! and will put himself on Mutae.....YEA!!!!

**Tim Zimmerlin (to Everyone)**: 2:01 PM: Also suggest: distributed algorithms view ala Paxos, quorums, Byzantine generals decision making, fault tolerance and recovery. Suggest Mesos/Kupernetes with Linux + Docker containers.

**Tim Zimmerlin (to Everyone)**: 2:03 PM: Finally suggest: federation view ala IdP, IdM, trust authorities, IPsec, ISAkey, certs, credentials, event messaging services.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:05 PM: Dear Tim: Good idea....also, our setup automatically records almost 99.999% of all content entered into the Chat Window....as a result, a Record of Comments is automatically generedted for Pw...without lifting a finger....go figure....? Also, we're sending along two (2) glossaries we believe are appropriate for the NBD-WG Security & Privacy Sub-WG.....That's our plan.....

**Tim Zimmerlin (to Everyone)**: 2:09 PM: Lexus-Nexus is a big data provider.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:11 PM: Direct from the Sedona Conference......best wishes, too.....:

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:11 PM: https://thesedonaconference.org/publications#ediscovery

**Tim Zimmerlin (to Everyone)**: 2:11 PM: Thomson Reuters Calais & Open Calais is a data discovery service.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:13 PM: Yes, Tim and I would like to discuss working on this.....with his permission.....of course.....

**gretchen gottlich (to Everyone)**: 2:13 PM: I will be happy to do a use case

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:13 PM: Correction; Tim, Gretchen & Pw would like to work on this, if they are agreeable.....?

**gretchen gottlich (to Everyone)**: 2:13 PM: yes

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:14 PM: Thank you Gretchen: 2/3rds of the Use Case Team is on-board......Hats off to us....

**gretchen gottlich (to Everyone)**: 2:14 PM: I will take the Twitter Analysis

**Korin Reid (to Everyone)**: 2:15 PM: i'll take healthcare

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:15 PM: Perfect: Pw loves Twitter Analysis.....too....thanks....

**gretchen gottlich (to Everyone)**: 2:16 PM: I will take the Twitter Use Case

**Russell (to Everyone)**: 2:17 PM: Im open to attempting some work on a healthcare, twitter, and #6

**gretchen gottlich (to Everyone)**: 2:17 PM: yes

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:18 PM: Dear Russell: Aren't there a couple of IU students building a Twitter Analysis start-up...?

**Russell (to Everyone)**: 2:18 PM: PW everyone is building twitter analysis startups

**Ann Racuya-Robbins (to Everyone)**: 2:19 PM: I would be interested in Human Face detection and either healthcare or data mining

**Russell (to Everyone)**: 2:19 PM: \* / are

**Korin Reid (to Everyone)**: 2:19 PM: I can do data mining in addition to healthcare

**Russell (to Everyone)**: 2:19 PM: Ann I could collaborate w you on the data mining

**Ann Racuya-Robbins (to Everyone)**: 2:20 PM: great

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:20 PM: Dear Russell: Yep, it's all part of FinTech, et....but these guys were in the news last year.....as in Forecasting stock-market fluctuations...sorta like this past weekends backdoor/under-the-table equity guarantees between soverign banks.....

**Russell (to Everyone)**: 2:20 PM: russell.reinsch@gmail

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:20 PM: Dear Russell: Here is our point of contact....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:21 PM: pwc.pwcarey@gmail.com

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:22 PM: Dear David:

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:22 PM: Should we send you a copy of the Chat Room after we're done today....?

**Tim Zimmerlin (to Everyone)**: 2:25 PM: Wo, first we need to size the six datasets. E.g., in AWS we could use S3 or a noSQL service.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:26 PM: Dear Sir: Point of Clarification....1). last week it was brought up that it might want to move forward without waiting for ISO to have a movement......and when we're satisified with our work....send it over the transome to the SITTING ISO folks....is that a correct and honest assessment....?

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:27 PM: The 'it' being this WG......for clarification.....it = NIST-NBD WG.....

**Tim Zimmerlin (to Everyone)**: 2:27 PM: Pw, today, Wo is leading the NBD-PWG full speed forward on our version 2.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:29 PM: Dear Tim: Ok....understood....thanks....we'll stop tossing out off-topic queries.....Respectfully yours, Pw

**Tim Zimmerlin (to Everyone)**: 2:29 PM: Pw, today, Jacob and Keith are covering ISO concerns and agendas.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:30 PM: Tim: Ok....Thanks....Pw

**gretchen gottlich (to Everyone)**: 2:30 PM: http://www.architectingforum.org/whitepapers/SAF\_WhitePaper\_2007\_4.pdf

**gretchen gottlich (to Everyone)**: 2:31 PM: sure

**gretchen gottlich (to Everyone)**: 2:31 PM: https://standards.ieee.org/develop/project/1723.html

**gretchen gottlich (to Everyone)**: 2:31 PM: https://standards.ieee.org/develop/project/1723.2.html

**gretchen gottlich (to Everyone)**: 2:32 PM: IEEE reference architecture foundation

**gretchen gottlich (to Everyone)**: 2:32 PM: above

**Tim Zimmerlin (to Everyone)**: 2:32 PM: Gretchen, I cannot open your links.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:33 PM: Dear Gretchen: How do we get a copy of this IEEE doc....?

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:34 PM: Please Note: We're not a member of IEEE....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:36 PM: Dear David: Does this diagram change with the Industry involved.....?

**gretchen gottlich (to Everyone)**: 2:38 PM: I have been in the Big Data ecosystem since 2011. what I do

**gretchen gottlich (to Everyone)**: 2:38 PM: https://www.linkedin.com/in/gretchengottlich

**gretchen gottlich (to Everyone)**: 2:38 PM: No have to go

**Robert Reyling (to Everyone)**: 2:38 PM: Can we put out a brief guide of what information are we expected to provide in our review of each use case (as assigned, etc)?

**gretchen gottlich (to Everyone)**: 2:38 PM: I am here to, to hopefully make sure that silly stuff doesn't happen

**gretchen gottlich (to Everyone)**: 2:39 PM: I am also working on being a patent agent

**Korin Reid (to Everyone)**: 2:39 PM: korin.reid@mckesson.com

**Robert Reyling (to Everyone)**: 2:45 PM: robert.reyling@wyle.com

**Russell (to Everyone)**: 2:46 PM: ciao

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:46 PM: Nope....good meeting....Sir......

**Robert Reyling (to Everyone)**: 2:46 PM: Good ideaon recurring introductions!