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

**NBD-PWD-2015/M0441**

**Source: NBD-PWG**

**Status: Draft**

**Title: Web chat and note from Meeting of Sept. 15, 2015**

**Action Items:**

1. Recapture: Walk through the selected use case using the NIST RA conceptual model:

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:

(please let me know if I have missed others)

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

**Chat Log D:\\_wo\1DMG\2015\\_BigDataWG\Docs\ChatLog 2015\_09\_15 15\_15.rtf**

**gretchen gottlich (to Everyone)**: 1:09 PM: My apologies. I have not had the time to map the Twitter Use Case to the Architecture document

**Mark Underwood (to Everyone)**: 1:12 PM: She is referencing Apache Marquet

**Korin Reid (to Everyone)**: 1:16 PM: parquet

**Mark Underwood (to Everyone)**: 1:16 PM: Oops :)

**Mark Underwood (to Everyone)**: 1:16 PM: My typo

**Mark Underwood (to Everyone)**: 1:17 PM: Korin -which biz group at McKesson are you with?

**Korin Reid (to Everyone)**: 1:18 PM: Mckesson Health Services

**Mark Underwood (to Everyone)**: 1:20 PM: Korin - got it

**Mark Underwood (to Everyone)**: 1:20 PM: Very helpful

**Korin Reid (to Everyone)**: 1:21 PM: My apology Mckesson Health Solutions

**Mark Underwood (to Everyone)**: 1:21 PM: ack'd

**Mark Underwood (to Everyone)**: 1:23 PM: What I'm seeing is Tim's screen

**Russell (to Everyone)**: 1:27 PM: David: thanks for that separation

**Russell (to Everyone)**: 1:28 PM: \*clarification on the two views

**Tim Zimmerlin (to Everyone)**: 1:34 PM: M399 v3

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:36 PM: Sorry for being late......Dear Sir: Just submitted our comments regarding a NIST DRAFT for Mobile IT Healthcare Cyber-security.....all five documents....were....'friendly'......unfortunately from our perspective friendly morfed into 5 Sales & Marketing Standards....with an application named Puppet....heading the list....In our personal and professional opinion.....we should be intelligent enough to come up with generic examples, without providing the vendors with a marketing tool....which they will use...eventually.....ethics aside....just our opinion.....Good meeting, too....

**Tim Zimmerlin (to Everyone)**: 1:40 PM: I suggest these volunteers start by identifying which big data analytic technique is optimal for their chosen NIST use case.

**Mark Underwood (to Everyone)**: 1:40 PM: PW - are you referring to NIST 1800-1 http://1.usa.gov/1JamwZR

**Mark Underwood (to Everyone)**: 1:41 PM: Gretchen - Nobody has to apologize for doing reimbursable work.

**Tim Zimmerlin (to Everyone)**: 1:42 PM: E.g., linear regression, clustering, neural network learning, etc. E.g., supervised, unsupervised, feedback.

**Russell (to Everyone)**: 1:43 PM: It wasnt me that time

**gretchen gottlich (to Everyone)**: 1:43 PM: If you could post that on the site or just e-mail me the document template I certainly follow that template

**gretchen gottlich (to Everyone)**: 1:43 PM: that was me typing. My bad my apologies.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:44 PM: Dear Mark: Yes.....Ver. a, b, c, d & e......here is the how to guide they're asking for comments due on the 25th of Sept.......this is an example of a NIST WG morfing over into a Sales & Marketing Publication for ....Puppet....go figure....:http://nccoe.nist.gov/sites/default/files/nccoe/NIST\_SP1800-1c\_Draft\_HIT\_Mobile-HowTo\_0.pdf

**Tim Zimmerlin (to Everyone)**: 1:44 PM: E.g., batch or streaming, once through or user interactive, etc.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:46 PM: Gretchen: Can you send me your email address so we can exchange buckets of knowledge regarding Twitters Stock Crash...and Use Cases for FinTech.....Go Team Twitter....

**Mark Underwood (to Everyone)**: 1:48 PM: PW - Yes, I received the call for comment earlier this summer via Mitre. The doc doesn't really address Big Data SnP

**Tim Zimmerlin (to Everyone)**: 1:48 PM: Actually, start by identifying "who" is analyzing the big dataset & "why" & the definition of "success", for all six NIST sample use cases.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:50 PM: Dear Tim: Shouldn't we back-up and select a specific industry for our Twitter Use Cases....as in some are highly regulated and some are loose cannons......Please respond......when you have a moment....

**gretchen gottlich (to Everyone)**: 1:50 PM: good call mark

**Tim Zimmerlin (to Everyone)**: 1:50 PM: sysML is a UML subset language.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:51 PM: Dear Mark: We agree.....it deserves consideration.....

**Tim Zimmerlin (to Everyone)**: 1:51 PM: HIPO, data flow, data structure...

**gretchen gottlich (to Everyone)**: 1:52 PM: visio is always a great back up ;)

**sharafudheen(Sharaf) (to Everyone)**: 1:52 PM: I totally agree with Wo that the current archetecture is very generic. I support mark and frank on not following a workflow approach

**Mark Underwood (to Everyone)**: 1:53 PM: As Frank said not UML but SysML or BPMN or others Frank mentioned

**sharafudheen(Sharaf) (to Everyone)**: 1:53 PM: May be we can have a high level diagram and then we can have splitup on each big blocks??

**Tim Zimmerlin (to Everyone)**: 1:54 PM: Pw, for the Twitter sample use case, you still need to imagine who, why, success, batch vs streaming, once through (e.g., detection) vs user interactive.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:55 PM: Dear Tim: True......also, reviewing the OSI model for incorporation into our Twitter Use Case can be beneficial....:http://www.proprofs.com/mwiki/index.php/The\_OSI\_Model

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:56 PM: Dear Tim: Beneficial in breaking out the complexities of the cyber-security issues associated at each phase of the data flow.....don't you agree....ummmmm ?

**Mark Underwood (to Everyone)**: 1:56 PM: FYI https://en.wikipedia.org/wiki/Federal\_enterprise\_architecture

**Tim Zimmerlin (to Everyone)**: 1:57 PM: Pw, agree fully.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:57 PM: Dear Tim: Lets go have lunch.......Application All 7

Presentation People 6

Session Seem 5

Transport To 4

Network Need 3

Data Link Data 2

Physical Processing 1

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:58 PM: Dear Gretchen: Pw's email address is.....pwc.pwcarey@gmail.com......Thanks...

**Tim Zimmerlin (to Everyone)**: 1:59 PM: First implement on NIST server cluster. Second implement on GENI, CloudLab, and Chameleon. Third implement on Amazon AWS. Fourth implement on Micrososft Azure. Fifth implement on Google Compute Engine.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:00 PM: Thanks Mark.....we'll review.....

**Cavan Capps (to Everyone)**: 2:02 PM: sorry am late and missed last week due to annual leave

**Cavan Capps (to Everyone)**: 2:03 PM: can we put in a draft use case? (one that has not been reviewed by the agency?)

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:04 PM: Dear Cavan: Last weeks meeting minutes are available for review at your convenience......

**Cavan Capps (to Everyone)**: 2:05 PM: thx

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:10 PM: Regarding Use Case strategy for Twitter.....GRC and Individual Security & Privacy are not on the list from the vendors perspective, rather they want to churn dollars as quickly as possible via Analytics using Twitter....as iin the following example: https://community.watsonanalytics.com/watson-analytics-use-case-3-ways-to-get-insights-from-twitter-data/

**sharafudheen(Sharaf) (to Everyone)**: 2:11 PM: where is the version 2 docs currently located?

**Russell (to Everyone)**: 2:14 PM: Sharaf - My understanding is the current versions are in Output Listing

**Russell (to Everyone)**: 2:14 PM: but

**Russell (to Everyone)**: 2:14 PM: Tim has said to me that Input Listings is the better folder

**Russell (to Everyone)**: 2:14 PM: for...

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:16 PM: Dear Sir: there are multiple Use Case templates; functional, regulatory, procedurial, technical, et al......depending upon the audience and the purpose and the industry......amazing, isn't it.....the more chefs will only spoil the broth....in our opinion......we should be able to determine what's good and what isn't......as in a real stinker.....so to speak.....in our personal and professional opinion....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:18 PM: Dear Folks: this is the most eclettic, organized, agenda lacking, knowledgeable, and friendly professional collection of folks we're been associated with......outside this WG....not so sure.....

**Russell (to Everyone)**: 2:19 PM: Wo: Ann has joined the meeting

**David W. Boyd (to Everyone)**: 2:22 PM: Geoffrey: Have you seen this book: Specifying Big Data Benchmarks - Is is from springer and is selected papers from the WBDB 2012 workshops.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:24 PM: David: Here is their link....http://link.springer.com/book/10.1007%2F978-3-642-53974-9

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:27 PM: Ann's name is Ms. Incognito.....no?

**Mark Underwood (to Everyone)**: 2:27 PM: Tim - It goes to a public Google XLS, so that is a possible issue / limitation with the Google Form approach. Though individual agencies could work around it by requesting a paper version (which we don't have yet :))

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:28 PM: Dear Tim: We don't trust Google as far as we can spit.....oopppps how did that slip out....

**Tim Zimmerlin (to Everyone)**: 2:30 PM: Frank, you are seeing the data oceans (e.g., GIG), data lakes, and data pools.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:30 PM: Dear Frank: Good hear.....we support you.....the Internet of Things is not designed for SNP from the ground up....rather it's another 'bolt on' solution.....go figure....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:31 PM: Good to hear Frank....good to hear....keep at it...Respectfully yours, Pw

**Tim Zimmerlin (to Everyone)**: 2:31 PM: Frank, I view the common ground as the set of all cyber - physical - social IT systems.

**Tim Zimmerlin (to Everyone)**: 2:32 PM: Only one other person agrees with me so far. Most restrict their focus to cyber - physical or cyber - social systems and concerns.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:33 PM: The smart phone platform is taking over any and all functions of commerce....end-to-end.....

**Tim Zimmerlin (to Everyone)**: 2:34 PM: Pw, do not tell anyone but that is NOT a phone…it is a powerful computing platform with many real time sensors.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:34 PM: Oh my gaud......

**Tim Zimmerlin (to Everyone)**: 2:34 PM: heh-heh

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:36 PM: Dear Frank: We thought Flowers and Tuttle built the first programmable computer during WWII....cracking the Lorenz cryptographic teletypewriter....?

**Tim Zimmerlin (to Everyone)**: 2:36 PM: The Open Group defines its mission as evolving the "nexus" of cloud computing, big data, mobile devices, and social networking.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:37 PM: So, we have an entire group supporting our opinions.....eh....Thanks Tim....

**Tim Zimmerlin (to Everyone)**: 2:37 PM: Pw, you would enjoy researching the Open Group, history, players, money, initiatives, etc.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:38 PM: Dear Tim: Will do....and isn't social networking a 'concept' rather than a technology.....so, can we kick it off the stage of technologies.....please?

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:41 PM: Dear Frank: When will they build Security & Privacy into the new technologies from the ground up....from day one.....?

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:42 PM: Out of those numbers......how much is useful, important data, rather than social entertainment....?

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:43 PM: Important....as in Mission Critical....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:47 PM: Come up with some useful algorithms....so venders can talk with each other.....via our Use Cases....something useful....something useable....how would you take in the data...data sets....--> algorithms--> architectures-->outcomes.......via Big Data....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:48 PM: How do you create your work environment.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:48 PM: Your own work environment....that is....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:49 PM: We have established a minimum baseline for Individual Security and Privacy haven't we....something they can use as a guide...?

**Ann Racuya-Robbins (to Everyone)**: 2:50 PM: I think Tim's suggestion has a lot of merit.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:51 PM: Well, if Ann thinks it's a good idea....we do too.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:55 PM: Nice meeting Sir......and Gretchen, too.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 3:11 PM: Whoever drives the meeting can number the agenda....7 minutes for discussion...then move on....table...or add to the next agenda for further discussion.....then lock-step down the agenda....ending with any new items....agreed....agreed....and close out the 45 minute meeting....and don't forget to stick to the time limits and provide the minutes to each attendee......End of story....Respectfullly yours, Pw

**gretchen gottlich (to Everyone)**: 3:15 PM: good night John boy