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

**NBD-PWD-2015/M0452**

**Source: NBD-PWG**

**Status: Draft**

**Title: Web chat from Meeting of Oct. 6, 2015**

**Chat Log D:\\_wo\1DMG\2015\\_BigDataWG\Docs\ChatLog 2015\_10\_06 15\_15.rtf**

**Pw Carey, Compliance Partners, LLC (Private)**: 12:58 PM: We'll try our microphone otherwise we'lll just converse via the Chat Window.....

**Wo Chang (to Pw Carey, Compliance Partners, LLC)**: 1:00 PM: Sounds good, thanks PW!

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:03 PM: Pw is putting himself on Mutae.....thank god....but looks forward to discussing Agenda Item #4....Twitter Use Case et al...

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:04 PM: Yes....we can hear you perfectly.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:06 PM: Background info. Re: Agenda Item #4:

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:06 PM: Block Chain\_Twitter & FinTech Oct. 6th, 2015

Blockchain is one solution....for anonomity, privacy & security...

3-Tier....

Networks/Carrier--->

FinTech Vendor--->

End User (Banks/Investment Banks...et)

MCKINSEY has warned banks that fintech startups are a threat to traditional banking’s profits, giving them a choice: fight them or join them. London’s fintech hubs are booming, with over 12,000 startups

TWITTER: It is set to make current interim boss and co-founder Jack Dorsey its

permanent chief executive, according to tech website Re/Code.

Dorsey was previously considered out of the running for the role due to his commitment to his other job as chief executive of

mobile payment company Square, which recently filed for an initial public offering.

The Twitter Government and Elections Handbook covers everything in minute detail, taking an impressive 25 pages just to reach the perhaps rather key question of: What is

Twitter? Twitter’s Bridget Coyne told NPR that the manual had b

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:08 PM: The Twitter Government and Elections Handbook covers everything in minute detail, taking an impressive 25 pages just to reach the perhaps rather key question of: What is

Twitter? Twitter’s Bridget Coyne told NPR that the manual had been “wildly popular”.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:09 PM: It's a good thing to off-load tasks to free up mind-experiments....and thinking time...too

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:11 PM: Fifteen Attendees and growing.....on the call today....

**Bob Marcus (to Everyone)**: 1:16 PM: Interfacing CPS, Big Data, and Cloud Models Slides will be at document 450 in NIST Big Data PWG input list. Not downloadable yet.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:17 PM: How to introduce our WG's Reference Architecure to the other groups and assist them as we appear to be further along on the steep learning curve...regarding Big Data/Cloud...even Mobile Eco-systems.....(open for corrections)....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:18 PM: The slides are on the database...Item No. 450....almost ready for downloading....will be shortly....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:20 PM: Considerations for the IoT Internet of Things, Cloud, Big Data et al......areas the other groups are looking at.......Now, lets consider Data Analytics.....for further analysis within the Reference Architecture......(aka: Big Data Analytics....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:21 PM: How will our WG's Architecture scale....with these other constructs.....for a mutual understanding of what they're doing....?

**sharaf (to Everyone)**: 1:22 PM: How can we get the details on CPS-PWD? is there any open repository of documents? can anyone clarify, pls?

**Bob Marcus (to Everyone)**: 1:22 PM: Interfacing with CPS requires outputting commands to actuators as well as inputting data from sensors

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:22 PM: Tim has some comments regarding collabration with these other WG's.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:23 PM: Dear Bob: A review of the Stuxnet virus is a good background starting point...SCADA et al.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:24 PM: Dear Bob: Is this a valid point....?

**Russell Reinsch (to Everyone)**: 1:24 PM: Need to step away for 2 minutes

**Russell Reinsch (to Everyone)**: 1:27 PM: Returned

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:28 PM: Dear Russell: Welcome back....

**sharaf (to Everyone)**: 1:29 PM: I have completed a white paper (research survey style) on "archetecture elements of WSN, IoT and security of big data". I have submittted to publish it in IEEE explore. Once it is completed, I will share to the team.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:32 PM: Dear Sharaf: Thanks....looking forward to reviewing it......

**Ann Racuya-Robbins (to Everyone)**: 1:34 PM: I am interested in UC #6

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:35 PM: Dear Ann: This sounds like a nice volunteering offer to me.....?

**sharaf (to Everyone)**: 1:35 PM: As discussed in last meeting, I am interested in UC: "BIg Data for Security"

**sharaf (to Everyone)**: 1:35 PM: I am still trying to get the sample data set...

**sharaf (to Everyone)**: 1:36 PM: As I am working in BI and DW, I am also interested in UC#6 with HANA.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:36 PM: Dear Sharaf: Is there a specific Use Case for Big Data for Security.....?

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:37 PM: HANA relating to SAP.....yes...?

**sharaf (to Everyone)**: 1:37 PM: It is the typical splunk usage scenario. Collect logs and other traces of system and analyse the data to produce system security measures (proactive / reactive)

**Tim Zimmerlin (to Everyone)**: 1:37 PM: Implementation teams require logistic support: 1) an Internet SW repository available for download to any cloud , grid, or cluster. 2) build scripts for targets like Linux and Windows. 3) install scripts for targets. 4) runtime scripts.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:38 PM: Dear Sharaf: We're aware of Big Data being used for Fraud Detection....?

**sharaf (to Everyone)**: 1:38 PM: SAP HANA is their in-memory big data implementation similar to SPARK technology

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:39 PM: Thanks....we didn't view HANA in that light.....thanks again...

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:40 PM: Dear Sharaf: In banking....Big Data Analytics is being used for Fraud Detection and KYC Know Your Customer puposes.....

**Tim Zimmerlin (to Everyone)**: 1:40 PM: The implementation teams will benefit from sharing their source code & scripts widely. The teams will also benefit fromt their appls running on many diverse systems.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:41 PM: Dear Sharaf: If you can investigate PEN testing it should be a benefit for any SAP Implementation....given enough lead time....

**Tim Zimmerlin (to Everyone)**: 1:41 PM: Later, we could run Red teams and Blue teams to test SnP.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:42 PM: Dear Sharaf: There should be some good resources within this WG to offer some insights.....at least we hope so......

**sharaf (to Everyone)**: 1:42 PM: Tim, do you think we use the test cases already provided by implemetors such as cloudera?

**Tim Zimmerlin (to Everyone)**: 1:42 PM: The NBD-PWG could publish metrics on performance.

**Tim Zimmerlin (to Everyone)**: 1:43 PM: Sharaf, I do not rule anything out.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:43 PM: Good idea.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:43 PM: Dear Sir: How could you.....

**Mark Underwood (to Everyone)**: 1:45 PM: RE: Splunk - this s/b seen as a general design pattern for SIEM Big Data. I'm not sure, but I haven't seen any specific advances from Splunk the company over what competitors are doing with SIEM

**Mark Underwood (to Everyone)**: 1:46 PM: Wo - Put me down as interested in reviewing the Healthcare UC4 for SnP (even if just walkthrough)

**Russell Reinsch (to Everyone)**: 1:46 PM: I am interested in more general Definitions for Vol 1

**Russell Reinsch (to Everyone)**: 1:46 PM: a and d

**Ann Racuya-Robbins (to Everyone)**: 1:47 PM: IHaing audio problems have submitted some changes to definitions yesterda

**Tim Zimmerlin (to Everyone)**: 1:47 PM: Mark, I use Splunk as an example of open source SW. Also, I have briefed their founder and team so they get it.

**Russell Reinsch (to Everyone)**: 1:48 PM: I am interested in harmonizing the RA also

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:48 PM: Dear Sir: Ann is still with us......

**Ann Racuya-Robbins (to Everyone)**: 1:48 PM: Yes I am interested in working on value.

**Angie Williams (to Everyone)**: 1:48 PM: I am interested in general Definitions for Vol 1 specifically value of big data

**Ann Racuya-Robbins (to Everyone)**: 1:48 PM: My audio is working right sorry.

**Russell Reinsch (to Everyone)**: 1:49 PM: My personal feeling is that the RA could be the better doc with which to organize the taxonomy; or to split the current Mindmap into two or more maps

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:49 PM: Do please follow-up with your interests......

**Mark Underwood (to Everyone)**: 1:49 PM: Tim - that would have been an interesting meeting (Splunk)

**Tim Zimmerlin (to Everyone)**: 1:49 PM: Mark, they shared and I shared.

**Russell Reinsch (to Everyone)**: 1:50 PM: Cavan can I get a sidebar with you on our dataset selection for use case 6 please

**Mark Underwood (to Everyone)**: 1:50 PM: Tim :)

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:52 PM: Dear Folks: Lets take a look at Big Data Blockchains.....via IBM, Mr. Softee, et all......regarding it's impact upon Big Data Privacy and Security (aka: SnP).....

**Cavan Paul Capps (CENSUS/ADRM FED) (to Everyone)**: 1:53 PM: sure, it won't be a dataset so much as a set of functions with statistical data collection, processing and release - if that makes sense

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:53 PM: Or....just the functional features of using Blockchains for SnP.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:54 PM: Sounds like an aggressive program....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 1:55 PM: Security + Privacy + Anonomity.....= Good news for the privacy of individuals....No...?

**Russell Reinsch (to Everyone)**: 1:57 PM: Russell for 6d

**Mark Underwood (to Everyone)**: 1:58 PM: The Vol 4 "template" is shorthand for an improved framework of the security fabric - E.g., it cites the NIST CIICF Oct 2013 proposal in the footnote on p.4. The same page refers to NIST 8062.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:01 PM: In support of Tim: Implementation.....leads us into the amazing world of Configuration best practices....including testing....et al....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:05 PM: The data has a life-cycle from beginning to expirery.....(aka: a data funeral of days hours and or weeks....perhaps.....)

**Cavan Paul Capps (CENSUS/ADRM FED) (to Everyone)**: 2:08 PM: Linux Rhel 7, on a reseach DMZ

**Cavan Paul Capps (CENSUS/ADRM FED) (to Everyone)**: 2:09 PM: 364gb RAM, 24 terabytes, 2 socket 16 cores

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:10 PM: Dear Cavan: Can we look that up.....(aka: Research DMZ....)....?

**Mark Underwood (to Everyone)**: 2:11 PM: Tim, maybe something could be worked out with Gnip, which Twitter recently acquired.

**Cavan Paul Capps (CENSUS/ADRM FED) (to Everyone)**: 2:11 PM: sorry, it is a Census site - thinking about a place to set up a proof of concept, I would have to verify that it could be used for a proof of concept

**Mark Underwood (to Everyone)**: 2:13 PM: Dave - roles & activities central for SnP as well

**Tim Zimmerlin (to Everyone)**: 2:24 PM: Bingo!

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:24 PM: How come Pw thinks FedRAMP is nothing but a big fat bag of nits......as in....nit, nit, nit.....Please Note: Pw takes no credit for his opinions.....here, now or anywhere else....Respectfully yours, Pw

**Tim Zimmerlin (to Everyone)**: 2:25 PM: Pw, have met and listened to the FedRAMP mgt. They are serious folks who understand things.

**Gregory Weidman (to Everyone)**: 2:26 PM: Thanks.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:27 PM: Dear Tim: Correct....we've met with those folks and read their documents from top to bottom and found loop-holes up the wazoo....no offense.....nice folks they maybe....the end result remains nitified.....Respectfullly yours, Pw

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:27 PM: Dear Tim: Lets do lunch sometime.....

**Tim Zimmerlin (to Everyone)**: 2:27 PM: Pw, how about FedRAMP plus NIST best practices?

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:28 PM: Dear Tim: That sounds much better as long as they listen to NIST Best Practices....Thanks for the work-around-solution....Pw

**Cavan Paul Capps (CENSUS/ADRM FED) (to Everyone)**: 2:28 PM: I am working with the interfaces, there are significant problems - particularly with security

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:29 PM: Thank you.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:32 PM: Here is our approach for our Twitter Real-time Analysis Use Case Re: FinTech....Open for Comments.....Block Chain\_Twitter & FinTech Oct. 6th, 2015

Our approach to this Use Case was to stress the Regulatory Obligations via a Vendor neutral approach....by not specifying (aka: naming) a particular vendors product...

Blockchain is one solution....for anonomity, privacy & security...

3-Tiers That have to work together rather seamlessly for encryption, anonomity, privacy and security reasons....as each FinTech transaction must be reconciled at the end of the day.....

Networks/Carrier--->

FinTech Vendor--->

End User (Banks/Investment Banks...et)

MCKINSEY has warned banks that fintech startups are a threat to traditional banking’s profits, giving them a choice: fight them or join them. London’s fintech hubs are booming, with over 12,000 startups

TWITTER: It is set to make current interim boss and co-founder Jack Dorsey its

permanent chief executive, according to tech website Re/Code.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:32 PM: Dear Sir: Please bring it up...as we don't have a microphone....sorry

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:34 PM: We think we must dial in via Skype but leave GoToMeeting to accomplish.....wish us luck.....

**Ann Racuya-Robbins (to Everyone)**: 2:34 PM: its a bit small

**Cavan Paul Capps (CENSUS/ADRM FED) (Private)**: 2:37 PM: could you hear me earlier when I tried to talk?

**Mark Underwood (to Everyone)**: 2:45 PM: Russell- Maybe we should have a separate chat about how to represent these things in a quasi-workflow scheme. It may not be "simpler" but it would be more canonical

**David W. Boyd (to Everyone)**: 2:52 PM: How about a Google Form like what Mark did for S&P

**Mark Underwood (to Everyone)**: 2:53 PM: Dave - indeed that would help conformity

**David W. Boyd (to Everyone)**: 2:54 PM: That is what we need a Pro-forma conformance template

**David W. Boyd (to Everyone)**: 2:55 PM: I think much of what is on the screen is too low a level to test an RA - this is getting down into software design.

**Mark Underwood (to Everyone)**: 2:55 PM: Wo wants to pour the recipe into Puppet

**Mark Underwood (to Everyone)**: 2:55 PM: Dave - Yes, some "abstract out" needs to happen

**Tim Zimmerlin (to Everyone)**: 2:55 PM: Puppet is good enough and getting better.

**David W. Boyd (to Everyone)**: 2:56 PM: So for puppet - we really need to just define the functional components and where they are deployed.

**David W. Boyd (to Everyone)**: 2:56 PM: But puppet an application would need to be written to do the data formatting described here.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:56 PM: Sorry folks: our mike doesn't work....but here's our Use Case strategy.....please comment....if the hardcopy can be brought up......Thanks....Pw......Block Chain\_Twitter & FinTech Oct. 6th, 2015

Our approach to this Use Case was to stress the Regulatory Obligations via a Vendor neutral approach....by not specifying (aka: naming) a particular vendors product...

Blockchain is one solution....for anonomity, privacy & security...

3-Tiers That have to work together rather seamlessly for encryption, anonomity, privacy and security reasons....as each FinTech transaction must be reconciled at the end of the day.....

Networks/Carrier--->

FinTech Vendor--->

End User (Banks/Investment Banks...et)

MCKINSEY has warned banks that fintech startups are a threat to traditional banking’s profits, giving them a choice: fight them or join them. London’s fintech hubs are booming, with over 12,000 startups

TWITTER: It is set to make current interim boss and co-founder Jack Dorsey its

permanent chief exec

**Tim Zimmerlin (to Everyone)**: 2:58 PM: To scale up to 100,000s of nodes, new resource orchestrators are required. It is wise to understand the limits of scaling. Minimize handwritten scripts.

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:58 PM: Twitter is a massive messaging service....and a massive tracking service......w/o privacy.....

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 2:59 PM: Dear Russell: During our research.....we bumped into a blog dealing exactly with your Use Case.....we'll send it along to you.....utilizing HDFS, HIVE.....filters.....et cetera....with Twitter Code included to test against......pwc.pwcarey@gmail.com

**Mark Underwood (to Everyone)**: 3:00 PM: Russell- Interesting point - I have taken "viz" to mean a placeholder for big data-enabled HCI innovations, We haven't really addressed it

**Cavan Paul Capps (CENSUS/ADRM FED) (to Everyone)**: 3:01 PM: Are we only diagraming at the functional level? Not talking about how anything is implimented?

**Mark Underwood (to Everyone)**: 3:01 PM: Viz can apply to Big Data ETL, not just at the "right hand" analytics side

**Tim Zimmerlin (to Everyone)**: 3:03 PM: Everyone, AWS re:Invent 2015 streams talks Weds & Thurs at https://reinvent.awsevents.com/

**Mark Underwood (to Everyone)**: 3:06 PM: Viz is a subset of HCI - i.e., we need to think of Big Data HCI implications. No, that is not a simplifying concept.

**Tim Zimmerlin (to Everyone)**: 3:07 PM: Mark, I want the 3D UI in "Iron Man" movie.

**Robert Reyling (to Everyone)**: 3:08 PM: Wo,

**Mark Underwood (to Everyone)**: 3:08 PM: Tim Yes ! forward and backward rewind

**Robert Reyling (to Everyone)**: 3:08 PM: Wo, Do you plan to brief on test results next week?

**Cavan Paul Capps (CENSUS/ADRM FED) (Private)**: 3:08 PM: I would like participate in security and privacy

**Russell Reinsch (to Everyone)**: 3:09 PM: I can put up 3 of the cases, I would like to develop the informational section of the healthcare use case before posting. Also Wo -

**Pw Carey, Compliance Partners, LLC (to Everyone)**: 3:09 PM: We'll fix our microphone issues before the next meeting.....sorry about the lack of sound....but do loook at our Twitter Use Case.....Thanks Pw

**Russell Reinsch (Private)**: 3:10 PM: I dont have authorization to upload docs to the NIST site; unless I am not trying in the right location

**David W. Boyd (to Everyone)**: 3:11 PM: AMEN Kavin - no standards for credentials.

**Mark Underwood (to Everyone)**: 3:11 PM: It's a value-add prop for Cloudera

**David W. Boyd (to Everyone)**: 3:12 PM: S&P might be the one place we can define some standard I/Fs

**Mark Underwood (to Everyone)**: 3:12 PM: Dave - def a stds opportunity

**Mark Underwood (to Everyone)**: 3:15 PM: Cavan - hope you brief us on this outcome later - even if no results

**Cavan Paul Capps (CENSUS/ADRM FED) (Private)**: 3:15 PM: sounds good