**NBD-Requirements WG Minutes August 13, 2013**

**Agenda**

1. Current Status of Use Cases and Template (see http://bigdatawg.nist.gov/\_uploadfiles/M0105\_v1\_9285264375.docx)
2. Relation of Requirements to Use Case Templates
3. Proposed Structure of Final Report (see Below)
4. Network Implications of Big Data (left over from last agenda)

We discussed the first three items of agenda in a long active session recorded in chat log. We did not use web audio and the audio quality was more reliable. The main discussion involved following topics

* New use cases recorded in yellow below and end of section 2 of draft
* The mapping between section 2 and 3 of report i.e. from use cases to requirements. Here we modified the scope of 3.7. We noted that lifecycle management (3.6 3.7) included Data Veracity and Data Quality. With this change there was a clear mapping between use cases entries and architecture categories of section 3
* Bob Marcus urged us to prioritize requirements.
* We are working with a compressed timeline. We can consider improving analysis after September 27 report delivered
* There was a significant discussion of the overall process and how the different working groups relate. We note that the architecture group inputted into design of use case template and now the requirements group will generate input requirements into architecture

We identified immediate action items below

**Action items:**

1. Draft format/style for Section 2 (converting use case tabular format to text format – ½ to 2/3 page (Geoffrey) before next week
2. all use case contributors should submit their use cases in two weeks
3. Create a map of use cases to requirements (Geoffrey)
4. create Google doc for editing (Wo will create the draft final report doc)

**Revised Proposed Contents of Report (changes in Yellow)**

**Executive Summary.**

**1 Introduction.**

1.1 Objectives.

1.2 How This Report Was Produced.

1.3 Structure of This Report.

**2 Use Case Scenarios – Vertical Applications.**

2.1 Use Case Template.

2.2 Earth Science: Atmospheric Turbulence - Event Discovery and Predictive Analytics.

2.3 Healthcare: Pathology Imaging/digital pathology.

2.4 Healthcare: Genomic Measurements.

2.5 Industry: Cargo Shipping.

2.6 Earth Science: Radar Data Analysis for CReSIS.

2.7 Particle Physics: Analysis of LHC Large Hadron Collider Data: Discovery of Higgs particle.

2.8 Consumer Service: Netflix Movie Service.

2.9 Internet Service: Web Search.

Sensors data (Clemson U.)

Environment Bio-diversity data (Yuri)

Digital humanity (Yuri)

Earthquake science data (NASA, Geoffrey is contacting)

Geo data (JPL)

Astronomy data (A Szalay)

EHR data (??)

Social networking analysis (referral from Geoffrey)

Disaster Application (PW)

E-Commence and multimedia (from SecNPrivacy subgroup)

Big data for private or hybrid cloud (human genome, Geoffrey)

Government operations: Data.gov (Wo or seek others to contribute)

Finance (PW)

**3 Use Case Requirements**.

3.1 Data Source Requirements:

3.2 Transformation Requirements:

3.3 Resource Requirements.

3.4 Data Usage Requirements:

3.5 Security & Privacy Requirements:

3.6 Lifecycle Management Requirements:

3.7 System Management and Other Requirements:

**4 Service Abstractions (SAs). Possibly Drop Section 4**

4.1 What is an SA?.

4.2 Service Abstractions Objectives.

4.3 Service Abstractions Management.

4.4 Types of Service Abstractions.

4.4.1 Data Service Abstraction.

4.4.2 Transport Service Abstraction.

4.4.3 Usage Service Abstraction.

**5 Conclusions and Recommendations.**

**6 Reference.**

**7. Appendix A – Raw Use Case Data**

**Chat Log**

(10:59 AM) Karen G joined.

(11:04 AM) \_Cherry Tom\_(\_IEEE-SA\_) joined.

(11:16 AM) William Miller (MaCT USA) joined.

(11:18 AM) William Miller (MaCT USA): Wo is this Ref. Architecture diagram updatred and is it posted?

(11:19 AM) Pw Carey (Compliance Partners, LLC) joined.

(11:19 AM) Karen G: I agree with Geoffrey [paraphrasing] that the use cases would not dictate the implementation. And with Wo, that the architecture needs to be flexible to support the requirements.

(11:20 AM) Pw Carey (Compliance Partners, LLC): That is correct as we understand the world....

(11:20 AM) Pw Carey (Compliance Partners, LLC): We agree with everybody....

(11:23 AM) Pw Carey (Compliance Partners, LLC): Each of the following, has different requirements, processes, procedures, GRC/CIA obligations...: Community Cloud, Public Cloud, Private Cloud, Hybrid Cloud via IaaS (Infrastracture as a Service), PaaS (Platform as a Service) and Saas (Software as a Service)...(aka: Hypervisors et al......background info.....Respectfully yours, Pw

(11:25 AM) Pw Carey (Compliance Partners, LLC): We should address Mobility via BYOD 'Bring Your Own Device' into the workplace.....No?

(11:27 AM) Pw Carey (Compliance Partners, LLC): All of the 'Current Solutions' are technical in nature....which is open to admendment(s)....no?

(11:28 AM) Karen G: Geoffrey, I completely agree - there needs to be a clear mapping from the use case sections, to the sections / topics of the architecture.

(11:30 AM) Pw Carey (Compliance Partners, LLC): Do we address Data Availability & Data Integrity...?

(11:31 AM) Karen G: Availability and integrity - great point - that could also map to veracity / robustness.

(11:32 AM) Karen G: perhaps add a brief description in the use case template, describing what type of info to provide in each section of the template.

(11:32 AM) Karen G: Geoffrey, I agree - annotate.

(11:32 AM) Karen G: :)

(11:33 AM) Pw Carey (Compliance Partners, LLC): Annotations sound like a nice solution.....to us....

(11:33 AM) Pw Carey (Compliance Partners, LLC): Brief, brief, and brief annotations is a good choice.....

(11:36 AM) Karen G: astronomy

(11:36 AM) Pw Carey (Compliance Partners, LLC): Astronomy....yes...

(11:36 AM) Pw Carey (Compliance Partners, LLC): Smart crowd....eh?

(11:39 AM) William Miller (MaCT USA): The Cargo Shipping use case includes sesnsor (i.e. RFID, GPS, Sensors)

(11:40 AM) William Miller (MaCT USA): i am looking into disaster response

(11:41 AM) Karen G: It is challenging to manage a team of volunteers. :)

(11:41 AM) William Miller (MaCT USA): geoffrey what is your e-mail address

(11:42 AM) Pw Carey (Compliance Partners, LLC): Pw is working on DR/BC and we sent a DRAFT COPY to Arnod Roy at Fujitsu.....

(11:42 AM) Geoffrey Fox: gcf@iu.edu

(11:42 AM) Pw Carey (Compliance Partners, LLC): As we recall, Jessie James was famous....

(11:42 AM) Karen G: For the EHR data use case, I would suggest either data from a large health organization [provider or payor], OR a use case around public health.

(11:43 AM) Pw Carey (Compliance Partners, LLC): We'll send our BC/DR draft to Mr. Fox...

(11:44 AM) William Miller (MaCT USA) disconnected.

(11:44 AM) Karen G: Have you considered use cases from govt?

(11:44 AM) PavithraKenjige (PK Technology) joined.

(11:44 AM) Karen G: OK, tks.

(11:44 AM) William Miller (MaCT USA) joined.

(11:44 AM) William Miller (MaCT USA): http://www.uicds.us

(11:45 AM) William Miller (MaCT USA): this is the use caes that i am working on

(11:45 AM) PavithraKenjige (PK Technology)69 joined.

(11:45 AM) PavithraKenjige (PK Technology) disconnected.

(11:46 AM) Pw Carey (Compliance Partners, LLC): We agree there should be cross-pollination of knowledge sharing.....

(11:47 AM) Eugene Luster (DISA CTO/R2AD) joined.

(11:47 AM) William Miller (MaCT USA): NASA is a great use caes

(11:48 AM) PavithraKenjige (PK Technology)69 disconnected.

(11:49 AM) PavithraKenjige ( PK Technology) joined.

(11:50 AM) William Miller (MaCT USA): do we have information on if data sets are using data compression?

(11:50 AM) Pw Carey (Compliance Partners, LLC): Would that include Hurrican Sandy...which turned into a great big BC/DR muck up

(11:51 AM) William Miller (MaCT USA): food condition monitoring is a project that is going on in Canada to monitor food in stores across the country

(11:51 AM) Geoffrey Fox: i agree hurricane issues related but jpl is just earthquake

(11:51 AM) William Miller (MaCT USA): we do not have this type of program here

(11:51 AM) Pw Carey (Compliance Partners, LLC): Ok...thanks....

(11:52 AM) William Miller (MaCT USA): no monitoring of temperature conditions or use of anything to prevent biological exposure or food spoilage

(11:52 AM) William Miller (MaCT USA): definitely needs to be a use case

(11:52 AM) William Miller (MaCT USA): a huge problem

(11:53 AM) Geoffrey Fox: please do it!

(11:53 AM) William Miller (MaCT USA): many government file systems are confident or classifed

(11:54 AM) Pw Carey (Compliance Partners, LLC): Multi-media with a strong security flavour.....(aka: Digital Rights)....who actually owns the data....good question....and not one that comes with an easy answer....

(11:55 AM) William Miller (MaCT USA): governemnt wants private big data clouds not public big data clouds

(11:56 AM) Pw Carey (Compliance Partners, LLC): We should be able to use N/A or "At this point in time"...or for future development.....

(11:57 AM) Pw Carey (Compliance Partners, LLC): Is that the one we watched on Hong Kong TV?

(11:58 AM) \_Cherry Tom\_(\_IEEE-SA\_): anyone involved in data.gov?

(11:58 AM) Pw Carey (Compliance Partners, LLC): Just a slight security joke....

(12:02 PM) Karen G: Geoffrey, I have some ideas wrt health.

(12:02 PM) Karen G: Yikes - feedback. :)

(12:03 PM) Sanjay Mishra (Verizon) joined.

(12:03 PM) Pw Carey (Compliance Partners, LLC) disconnected.

(12:03 PM) Pw Carey (Compliance Partners, LLC) joined.

(12:04 PM) William Miller (MaCT USA) disconnected.

(12:04 PM) Karen G: Here's one list of high level industries:

(12:06 PM) Karen G: Wo, are you suggesting modifying the template from a tabular format, to formatted text?

(12:08 PM) Pw Carey (Compliance Partners, LLC): We'll get my entire team to look for a couple of 'Finance' Use Cases....if that's ok...?

(12:08 PM) Karen G: @ Pw Carey - that sounds great.

(12:09 PM) Pw Carey (Compliance Partners, LLC): Will do....

(12:09 PM) Pw Carey (Compliance Partners, LLC): We're a GRC Application Security Analyst, CISA, CISSP....Consultancy....

(12:10 PM) Pw Carey (Compliance Partners, LLC): Based in Barrington, IL USA

(12:10 PM) Bob Marcus(ET-Strategies): To help the other subgroups, it will be very valuable to extract specific high priority Requirements fro the Use Cases.

(12:11 PM) Karen G: Here's one view of industries; I think several of these are already represented in the use cases: [sorry for the formatting]: Aerospace Automotive Banking Chemicals and petroleum Communications Consumer products Defense EducationElectronicsEnergy and utilities Financial markets Government Healthcare Insurance Life sciences Media and entertainment Metals and mining Retail Travel and transportation

(12:12 PM) Karen G: I agree with Bob Marcus' comment immediately above.

(12:12 PM) Karen G: And, imo this could be a joint effort among the subgroups.

(12:12 PM) Karen G: OK, got that.

(12:13 PM) Pw Carey (Compliance Partners, LLC): So, Pw had volunteered to work on the FINANCE USE CASES...Thank you Pw....Respectfully yours, Pw...and Bob is right, too....

(12:14 PM) Geoffrey Fox: great

(12:14 PM) Karen G: Bob - that would require prioritizing across industry needs; I'm not sure that is feasible. Maybe possible to identify common needs across use cases.

(12:15 PM) Karen G: The other subgroups don't have time to read the use cases?

(12:15 PM) Karen G: So Bob, do you need more detail in specific areas?

(12:16 PM) Karen G: Generally speaking, we prioritize requirements based upon a specific organization;s needs.

(12:17 PM) William Miller (MaCT USA) joined.

(12:17 PM) Karen G: Is there a current, detailed, project plan - with deliverables and dates across the subgroups?

(12:17 PM) William Miller (MaCT USA) disconnected.

(12:19 PM) Karen G: Bob, great point - I do think that by having use cases that are input from volunteers, it is rather difficult to assign priority.

(12:19 PM) Geoffrey Fox: agree

(12:20 PM) \_Cherry Tom\_(\_IEEE-SA\_): Perhaps a matrix with use cases listed down and requirements across and do checks?

(12:20 PM) Karen G: So, most frequently cited requirements. Still, without a comprehensive / representative sample, it can't be 'scientific'.

(12:21 PM) Karen G: Industry 'A' might have a requirement for GPS data of \_\_\_\_\_\_\_\_ volume, and I have no need of GPS data. :)

(12:21 PM) William Miller (MaCT USA) joined.

(12:22 PM) Geoffrey Fox disconnected.

(12:22 PM) Karen G: @Cherry Tom - great idea. And Wo / Bob just discussing a cross reference.

(12:23 PM) Geoffrey Fox joined.

(12:25 PM) Karen G: Still, need to acknowledge that this is a sample, and not a representative sample.

(12:25 PM) Pw Carey (Compliance Partners, LLC): We'll send you our notes for the Public Chat .... and you can come back in.....

(12:25 PM) Karen G: Yes Wo, I am just suggesting that priorities will vary across industries.

(12:25 PM) Pw Carey (Compliance Partners, LLC): For Mr. Fox.....

(12:25 PM) William Miller (MaCT USA): gps data is inclusive of an application

(12:26 PM) Karen G: So, we can identify requirements; cannot provide a prioritized list.

(12:26 PM) Karen G: Much as we might like to do so.

(12:26 PM) Karen G: Not me :)

(12:27 PM) Karen G: Perhaps they can dial direct.

(12:27 PM) Pw Carey (Compliance Partners, LLC): We just hung-up our phone connection...

(12:32 PM) Pw Carey (Compliance Partners, LLC): We can hear the audio after a fashion....

(12:32 PM) Karen G: Is there a current, detailed, project plan - with deliverables and dates - across the subgroups? Seems that would be helpful now.

(12:32 PM) Karen G: Good discussion.

(12:33 PM) Karen G: I see.

(12:33 PM) Karen G: Wo, maybe the Definitions team input & google doc will help with drafting that.

(12:33 PM) PavithraKenjige ( PK Technology) disconnected.

(12:34 PM) PavithraKenjige ( PK Technology) joined.

(12:34 PM) Pw Carey (Compliance Partners, LLC): Hear, hear.....

(12:35 PM) Karen G: Wo, I'm actually thinking of the discussion on yesterday's call.

(12:35 PM) Pw Carey (Compliance Partners, LLC): Pw's email...gmail is...pwc.pwcarey@gmail.com

(12:36 PM) Pw Carey (Compliance Partners, LLC): Ok will do...what is your email address...?

(12:36 PM) Karen G: One other observation: If you are looking for additional use cases, I believe it would be best to reach out to specific folks. A broadcast request might not get much response.

(12:37 PM) Karen G: Right - Definition call - might help you with the executive summary & section 1 Introduction.

(12:37 PM) Karen G: OK.

(12:37 PM) Karen G: Sounds good.

(12:37 PM) \_Cherry Tom\_(\_IEEE-SA\_): was the joint subgroup call postponed from this week?

(12:39 PM) Karen G: Section 4 sounds like an architectural notion.

(12:40 PM) Karen G: Having a single document for Requirements, right?

(12:42 PM) Wo Chang (NIST): Karen: you could be very corect that Section 4 may belong RA.

(12:42 PM) Wo Chang (NIST): Karen: yes, it will be a single Google doc document.

(12:44 PM) Karen G: Yes, we have discussed that requirements will vary across industries.\

(12:44 PM) Karen G: And, imo, it isn't possible to prioritize reqts across industries / verticals.

(12:45 PM) PavithraKenjige ( PK Technology) disconnected.

(12:45 PM) Karen G: So, current speaker makes an excellent point.

(12:45 PM) PavithraKenjige (PK Technology) joined.

(12:47 PM) Karen G: The use cases are exemplars for a certain industry, and not necessarily a comprehensive view across ALL industries ... if that makes sense.

(12:47 PM) Geoffrey Fox: yes exemplars

(12:48 PM) Karen G: Wo, Health IT folks are already implementing 'big data' solutions around population health. Are these solutions generalizable to other industries? Some aspects are, and some aspects are not.\

(12:50 PM) Geoffrey Fox: why not ask those IT people for use case

(12:51 PM) Karen G: I think that PK's comment is very relevant. It gets to the definition of a use case. To me, a use case defines how an actor interacts with a 'system' to accomplish a goal. I believe that in this case, the notion of a use case is more around potential technologies that can support - as Wo just said - generalized needs.

(12:52 PM) \_Cherry Tom\_(\_IEEE-SA\_): these requirements are a starting point for verticals, not complete. A representative set of industry people could discuss them and determine gaps?

(12:53 PM) Karen G: Certainly security and privacy reqts vary by industry and by geo.

(12:55 PM) Karen G: Yes, have asked about the project schedule & deliverables.

(1:01 PM) Karen G: From this discussion - if the goal is to establish guidance that is generalizable across industries....

(1:03 PM) Karen G: ... then a review of existing documentation, and interfacing to similar work efforts, would help. For example, what are other organizations already doing in the health IT space? in tele? in finance? ...

(1:03 PM) William Miller (MaCT USA) disconnected.

(1:05 PM) Karen G: Excellent comment - what are the interfaces between the subgroups. OK, gotta run!

(1:06 PM) Karen G disconnected.

(1:10 PM) Pw Carey (Compliance Partners, LLC): Good meeting....and best wishes, too...

(1:10 PM) Pw Carey (Compliance Partners, LLC): Respectfully yours, Pw

(1:12 PM) Pw Carey (Compliance Partners, LLC): Tomorrow's meeting...Wednesday, Aug. 14th is postponed.....

(1:14 PM) \_Cherry Tom\_(\_IEEE-SA\_) disconnected.

(1:14 PM) PavithraKenjige (PK Technology)54 joined.

(1:14 PM) William Miller (MaCT USA) joined.

(1:14 PM) PavithraKenjige (PK Technology) disconnected.

(1:16 PM) Pw Carey (Compliance Partners, LLC): Just buy everybody lunch...will be fine....

(1:17 PM) PavithraKenjige (PK Technology)54 disconnected.

(1:18 PM) Pw Carey (Compliance Partners, LLC): Good meeting...thank you....

(1:18 PM) Pw Carey (Compliance Partners, LLC) disconnected.