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

**NBD-PWD-2016/M0541**

**Source: NBD-PWG**

**Status: Draft**

**Title: Web chat from Meeting of July 5, 2016**

**Chat Log D:\\_wo\1DMG\2015\\_BigDataWG\Docs\ChatLog 2016\_07\_05 15\_03.rtf**

**Russell C Reinsch (to Organizer(s) Only)**: 1:04 PM: Meeting notes; minutes: Cavan, Frank, Karl Stiller, Nancy, Russell, LeAnn [ ? ]

**Russell C Reinsch (to Organizer(s) Only)**: 1:05 PM: Tim,

**Russell C Reinsch (to Organizer(s) Only)**: 1:07 PM: Ann,

**Russell C Reinsch (to Organizer(s) Only)**: 1:11 PM: Frank: ISO wrong, recent training materials: if no shall clause, cannot appear as a normative reference. Real technical stuff.

**Russell C Reinsch (to Organizer(s) Only)**: 1:13 PM: Second line of an address field = optional field. Geoffrey in attendance.

**Russell C Reinsch (to Organizer(s) Only)**: 1:15 PM: Phone caller: not a lot of vetting happens at ISO?

**Russell C Reinsch (to Organizer(s) Only)**: 1:15 PM: Frank: they dont interpret docs during standard implementations.

**Russell C Reinsch (to Organizer(s) Only)**: 1:18 PM: Nancy: TOC in 1500-1. Proposing a re-org to reflect same structure as RA.

**Russell C Reinsch (to Organizer(s) Only)**: 1:19 PM: Concepts: 2.1, 2.2, 2.3 ~ overview.

**Russell C Reinsch (to Organizer(s) Only)**: 1:21 PM: Frank: Concepts are right [linguistic ], r12ps, Definitions is part of Concepts.

**Russell C Reinsch (to Organizer(s) Only)**: 1:26 PM: N: Part 3 = Engineering Concepts. Innovations, new in eng. Nosql, in-memo, r/t... containers?

**Russell C Reinsch (to Organizer(s) Only)**: 1:28 PM: F: Gartner bricks and patterns.

**Russell C Reinsch (to Organizer(s) Only)**: 1:29 PM: Eric Harper in attendance.

**Russell C Reinsch (to Organizer(s) Only)**: 1:33 PM: Definitions volume vs additional Foundations vol.

**Cavan (to Everyone)**: 1:35 PM: Since various "big data software solutions are moving faster than concensus definitions" perhaps we provide as was suggested different patterns of arch (hope I am using "patterns" correctly)

**Russell C Reinsch (to Organizer(s) Only)**: 1:38 PM: Moving Data Science to a section 4. Analytics lifecycle [not data lifecycle]; nothing on vis12n yet. Exploratory is hard now. How well models working [2], and explanatory have not changed much.

**Russell C Reinsch (to Organizer(s) Only)**: 1:40 PM: Metadata treated in 4.5. R12p to HPC, [4.7] compute intense vs. data intense. What is diff for application. How does A change because framework changed. Ana [sql [only]] was separate before; cannot be separated now.

**Russell C Reinsch (to Organizer(s) Only)**: 1:42 PM: A moved to the data. Section 4 = use of the data. Sec 3 = Framework.

**Russell C Reinsch (to Organizer(s) Only)**: 1:43 PM: Sec 5 = Ownership. One of fundamental concepts that has changed a lot.

**Russell C Reinsch (to Organizer(s) Only)**: 1:49 PM: 6 = Governance / Mgmt. Everybody in attendance has an input on this topic. Academic world struggling w IRBs.

**Tim Zimmerlin (to Everyone)**: 1:54 PM: Cavan, Apple sues Google...

**Cavan (to Everyone)**: 1:57 PM: @Tim interesting

**Ann Racuya-Robbins (to Everyone)**: 2:20 PM: I think that is a good idea Nancy.

**Cavan (to Everyone)**: 2:23 PM: Actually the issue of the value of data is a difficult topic. Most things that have subjective value have that value quantitied by the supply and demand of the market

**Russell C Reinsch (to Organizer(s) Only)**: 2:24 PM: 7. SnP. 8. Implications, interfaces.

**Ann Racuya-Robbins (to Everyone)**: 2:24 PM: Agree Nancy about a sense of the big picture

**Cavan (to Everyone)**: 2:24 PM: There is discussion, can the market really determine the value of poverty data etc.

**Russell C Reinsch (to Organizer(s) Only)**: 2:26 PM: RA can be technical. Vol 1 can be conceptual.

**Russell C Reinsch (to Organizer(s) Only)**: 2:26 PM: Nancy unable to post directly to Drive

**NANCY GRADY (to Everyone)**: 2:30 PM: There are five fundamental drivers that determine if you have a big data problem:

(1) Irregular or heterogeneous data structures, their navigation, query, and data typing (variety);

(2) The need for computation and storage parallelism and its management during processing of large data sets (volume);

(3) Descriptive data and self-inquiry about objects for real-time decision-making (validity/veracity);

(4) The rate of arrival of the data (velocity); or

(5) Presentation and aggregation of such data sets (visualization)

**Frank Farance (to Everyone)**: 2:30 PM: Big Data: convergence of problems in one or more: (1) irregular or heterogeneous data structures, their navigation, query, and datatyping (variety); (2) the need for computation and storage parallelism and its management during processing of large data sets (volume); (3) descriptive data and self-inquiry about objects for real-time decision-making (validity/veracity); (4) the rate of arrival of the data (velocity); or (5) presentation and aggregation of such data sets (visualization)

**Tim Zimmerlin (to Everyone)**: 2:32 PM: Big data is whatever a stakeholder says is their big data.

**NANCY GRADY (to Everyone)**: 2:32 PM: @Tim, we don't want that since no one will understand anyone else when they say they do big data

**Tim Zimmerlin (to Everyone)**: 2:33 PM: Nancy, I know that...

**Tim Zimmerlin (to Everyone)**: 2:34 PM: The world is in the midst of a wholesale transition to networked cloud computing...big data is a set of emergent concerns shared by stakeholders.

**Tim Zimmerlin (to Everyone)**: 2:35 PM: Frank's definition is good enough...it is important to progress beyond these definitions.

**Russell C Reinsch (to Organizer(s) Only)**: 2:57 PM: Focus the taxon on the RA. Our taxon 1/2 data 1/2 ... and what is the value of our taxon.

**Wo Chang (to Everyone)**: 2:58 PM: Nancy: I need to take a call at 3:00PM but you all could continue.

**Russell C Reinsch (to Organizer(s) Only)**: 3:00 PM: Frank: not seeing the rationale.

**Russell C Reinsch (to Organizer(s) Only)**: 3:01 PM: We could get more out of a concept map.

**REYLING, ROBERT A CTR USAF AFMC AFLCMC/HN (to Organizer(s) Only)**: 3:02 PM: Can you make me the presenter?

**REYLING, ROBERT A CTR USAF AFMC AFLCMC/HN (to Everyone)**: 3:02 PM: End of Meeting is 3:03 EDT

**Wo Chang (to REYLING, ROBERT A CTR USAF AFMC AFLCMC/HN)**: 3:02 PM: Sorry robert that I see you chat late!

**Wo Chang (to REYLING, ROBERT A CTR USAF AFMC AFLCMC/HN)**: 3:02 PM: we could do it next week.