## Introduction

The Reference Architecture (RA) subgroup met for the first time today. The first 20 minutes of the meeting were lost to technical issues. There followed a brief introduction to the proposed subgroup charter. Orit Levin, the subgroup lead, introduced herself and presented group goals. The floor was then opened to discussion, the highlights of which follow. Orit Levin presented her RA paper as submitted to the input documents (NIST BD WG doc M0017). This led to in-depth exploration and discussion of the elements in Orit’s Big Data Ecosystem RA, slide 3 of that document. Action items were assigned to the co-chairs and member body with concluding comments by Wo Chan, NIST. The meeting concluded 13:02 EST.

## Attendees

|  |  |
| --- | --- |
|  | **Name** |
| 1 | aaa |
| 2 | Alicia Zuniga-Alvarado/AZA |
| 3 | Arnab Roy (Fujitsu) |
| 4 | Bob Marcus (ET-Strategies) |
| 5 | Carl Buffington |
| 6 | Dave Raddatz (SGI) |
| 7 | Don Krapohl |
| 8 | Eugene Luster (DISA CTO/R2AD) |
| 9 | gary mazzaferro |
| 10 | Glenn Wasson\_SAIC |
| 11 | James Ketner |
| 12 | Jill Gemmill(Clemson U) |
| 13 | Mike Seablom |
| 14 | Nancy Grady (SAIC) |
| 15 | Orit Levin |
| 16 | Orit Levin - Moderator |
| 17 | PavithraKenjige |
| 18 | QUYEN NGUYEN (NARA) |
| 19 | Rupinder Singh |
| 20 | Scott Brim (Internet2) |
| 21 | Serge Manning (Huawei USA) |
| 22 | Shawn Miller VA |
| 23 | Tim Zimmerlin (Automation Technologies) |
| 24 | Vivek navale |
| 25 | William Miller (MaCT USA) |
| 26 | Wo Chang (Host, NIST) |
| 27 | Yuri Demchenko (UvA) |
|  |  |

## Minutes

The meeting followed the Agenda

1. Technical issues caused 20 minute delay to the start of the meeting
2. Don Krapohl gave the URL to the proposed charter and began a brief introduction of the charter contents and the co-chairs. No comments followed.
3. Orit Levin presented introduction to goals for the subgroup
   1. We (the whole NIST BD WG) must come up quickly with use cases, requirements, and taxonomies which must map to each other.
   2. By presenting existing architectures hope is that we can find useful examples already done.
4. Orit Levin opened the floor to suggestions about how we move forward
   1. Eugene Luster notes that everything we discuss is public so everything will we say must be based on public (non-private) information. Also asked question of what is the first thing we wish to put out, understanding that taxonomy and other groups are working in parallel.
      1. Orit Levin pointed to the charter noting the deliverables mentioned. The goal is to make a deliverable that is useful to the other subgroups.
   2. Vivek Navela, NARA, asked if we should broadly categorize domains? Each community has its own frameworks. Asserts that we should consider industry-specific frameworks and collect them.
   3. Eugene Luster pointed out that we should be looking for Big Data in use, not just Big Data in general. “It’s not just about cloud. Particle physics alone has tremendous uses now.”
   4. William Miller put up a link for GoToMeeting and the web demonstration moved to that platform.
   5. Bob Marcus noted that we need to stick to core RA as abstraction since we have an initial deadline of 27 Sep. We may be able to do domain-specific following.
   6. Joe Mazzaferro posted to Taxonomy reflector other RA examples and will report them to the RA reflector
   7. Vivek Navela suggested the use of Federal Enterprise Architecture (FEA) for structured approach to building the RA and gave an introductory explanation of FEA as an RA abstraction.
   8. Gary Mazzaferro indicates that as a framework for implementation FEA has drawbacks. Its flexibility is a double-edged sword requiring governance to focus it. Proposes we focus on specific stakeholders and contrive messaging to address those issues. From NIST Cloud he noted it became important that specific messages and stakeholders be identified.
5. Orit Levin provided a review of Big Data Ecosystem RA as submitted in her paper submitted as document M0017 (http://bigdatawg.nist.gov/\_uploadfiles/M0017\_v1\_7268262499.pdf).
   1. Tim Zimmerman asserts coherent and cohesive data flow is what makes big data work. Also noted that he is concerned about scope creep resulting from trying to define RA for all domains.
      1. Orit Levin indicated the RA group should seek a common language to communicate across different domains.
   2. Tim Zimmerman studied all three diagrams in the document and requested highlight of which facets are important to and differentiations of big data as contrasted to existing data architectures.
      1. Orit Levin: all relevant to big data; among the things new to Big Data are decoupling of old functions and outsourcing some/many of them (including using SaaS for generic functions, PaaS, and IaaS for storage approaches); all need to be presented, while their priorities and relevance to NIST should be evaluated after the initial RA is developed.
   3. Vivek Navela asked how the presented RA addresses data preservation. How will this information be accessible and contextually preserved over time?
      1. Orit Levin: This is unknown.
      2. In data transformation the contextual information is preserved by the infrastructure elements and is mutable in time.
   4. Gary Mazzaferro questioned Orit Levin slide 3 on RA asking where in the diagram the interfaces will exist between elements in the RA. Orit asserts these exist but most are proprietary. Horizontal lines in the diagram indicate cross-cutting interfaces that may be open source.
      1. Gary Mazzaferro responded noting the differences in definition of interface. Protocols and standards already exist in these interfaces.
   5. Gary Mazzaferro in the transformation block of the RA document, are the blocks standardization-based (in the aggregation, collection, data mining, matching section)?
      1. Orit Levin suspects that many are not but rather they are likely vertical-specific. There may be value in standardizing some of these.
      2. Also privacy and security concerns impact here as data may change hands in this part.
      3. Gary Mazzaferro asserts that RA with specific design elements should be examined closely. Privacy and security were specifically listed as elements that may belong to lower-level [sub-block] insertion points in the architecture.
   6. Jill Gemmill, Clemson requested in chat that Orit Levin talk some more about Big Data and individual data transfer? See chat.
      1. Orit Levin indicated her original thinking was BD requires specialized techniques. For example, is data exchanged by reference or by value?
6. Action Items
   1. Co-chairs need to get the presentation technology working
   2. Membership, submit and read different approaches to building an RA. Orit Levin requested for presenters of other approaches, 30 minute maximum.
   3. Wo Chan added we need to identify key components of a RA. No assignment given
7. Concluded 13:02 EST.

Respectfully submitted by Don Krapohl, 7/13/2013