



# NBD-WG: Thinking Forward

Wo Chang, NIST

## DISCLAIMER:

The following presentation is for suggestion only and is not a requirement for NBD-WG/Subgroups.

# Good Approach (outside-the-box):

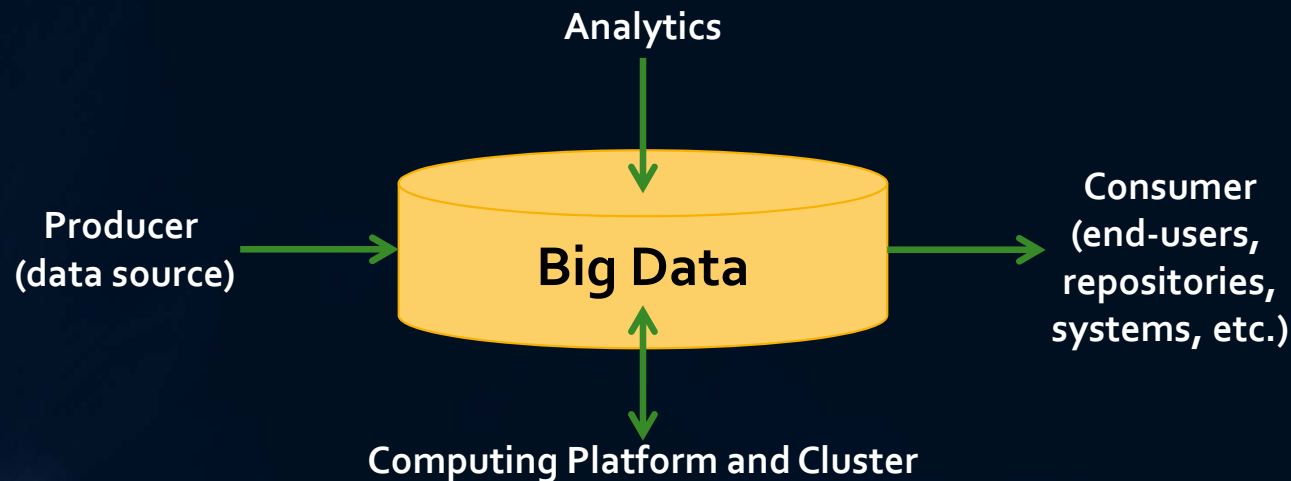
- Imagine the Web 20 years ago...  
...only a presentation tool with hyperlink using simple HTTP
- Web today and beyond  
Central ecosystem which allows value-added inside- and/or outside-the-box

Simple, flexible, and extensible!



# Big Data Current Approach (inside-the-box):

- How to optimize analytics tools
- How to better structure larger datasets
- Others...



They are indeed important!

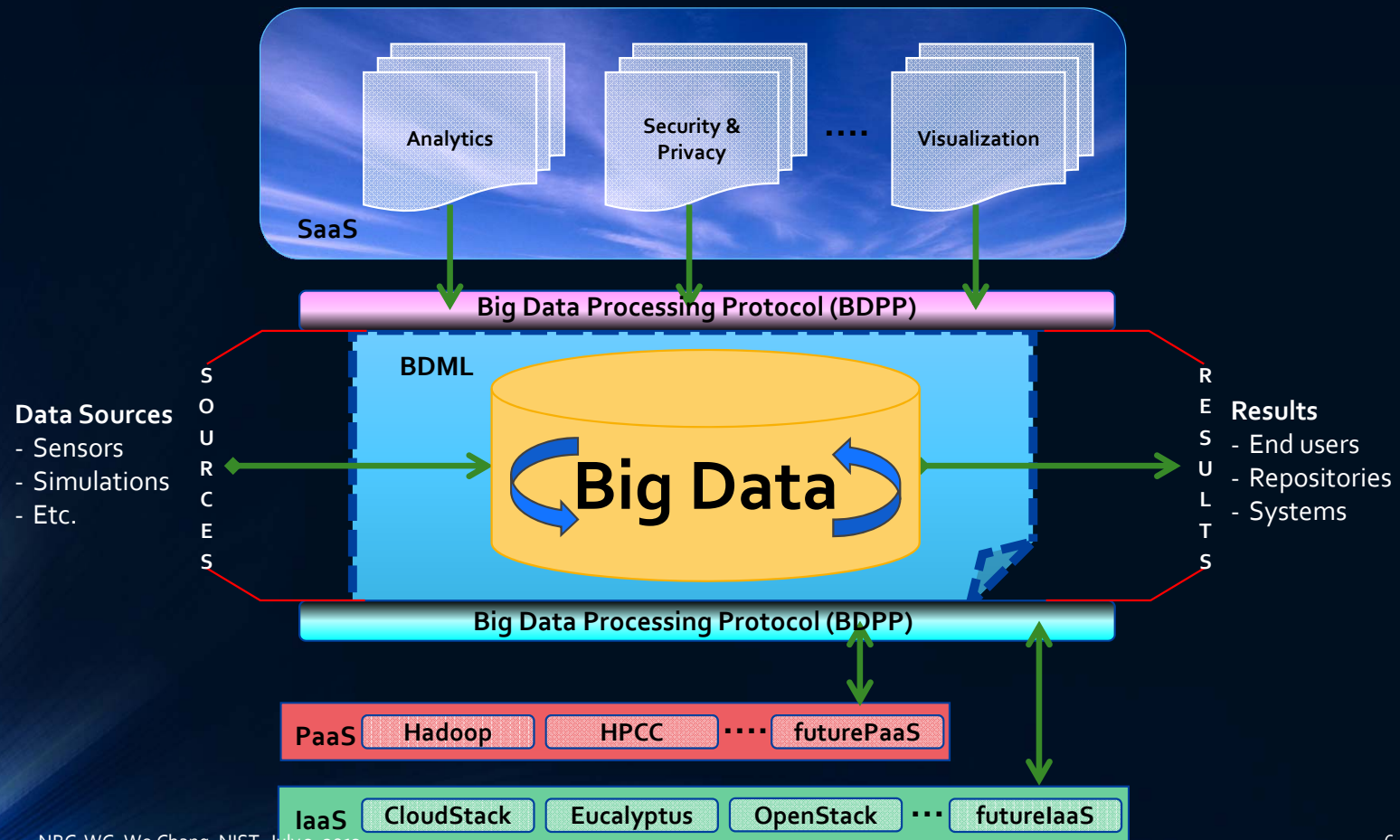
## Semi-new Approach (outside-the-box):

- How to orchestrate technology-agnostic analytics tools with secure mechanisms for end-users / systems consumption
- How to enable Big Data stakeholders to pick-and-choose technology-agnostic analytics tools for processing and visualization in any computing platform and cluster
- How to allow value-added from Big Data service providers

Simple, reusable, and extensible!



# Big Data Processing Protocol (BDPP) & Big Data Markup Language (BDML)



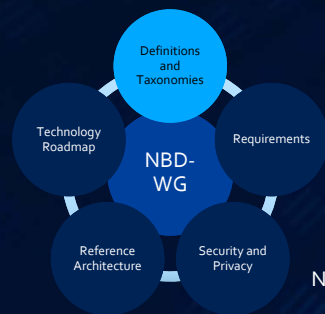
# SUBGROUPS

## AND THEIR TASKS





## Definitions and Taxonomies



NBG-WG, Wo Chang, NIST, July 3, 2013

### Base-line Definitions:

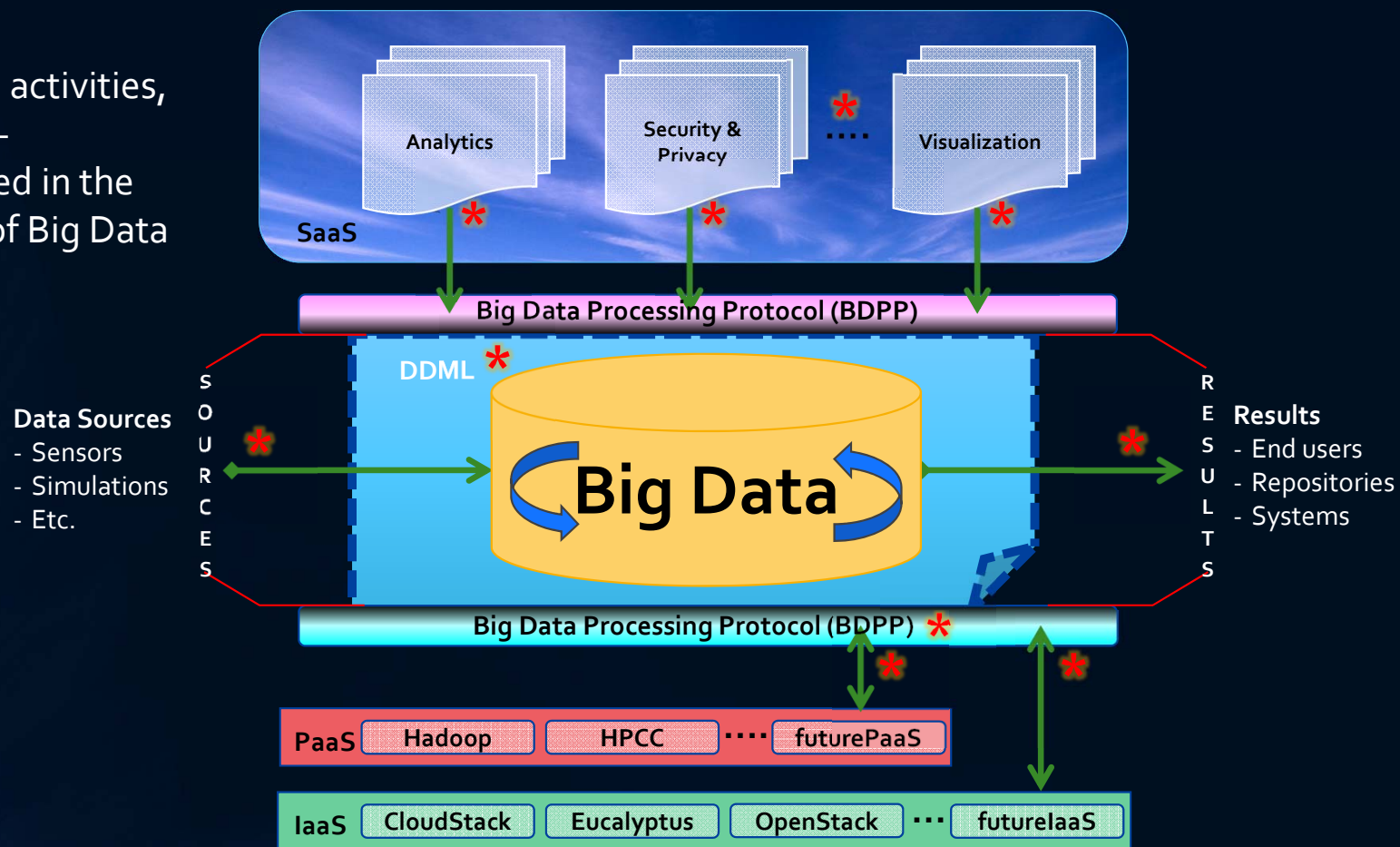
**Big Data** refers to digital data volume, velocity and / or variety [veracity] that:

- ✧ Enables novel approaches to frontier questions previously inaccessible or impractical using current or conventional methods; and / or
- ✧ Exceeds the capacity or capability of current or conventional methods and systems
- Factual information
- Discussion on Big Data characteristics
- We might want to extend what Big Data can do (or is it too early to say?)



## Taxonomies –

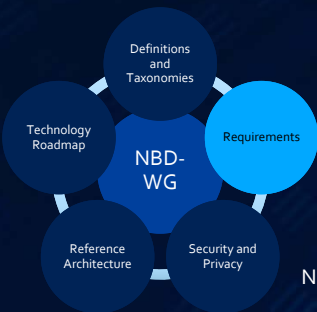
Explore actors (roles), activities, components, and sub-components associated in the surrounding (with \*) of Big Data





## Requirements

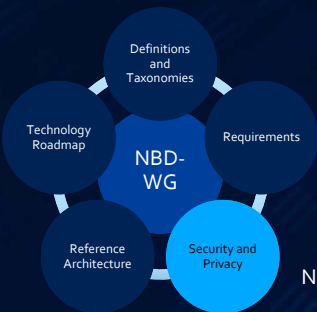
- Collect diversified use cases and identify their requirements
- Identify BDPP and BDML usage scenarios
- Identify what components are needed to enable BDPP and BDML



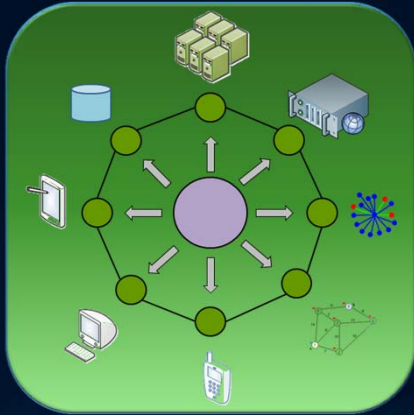


## Security and Privacy

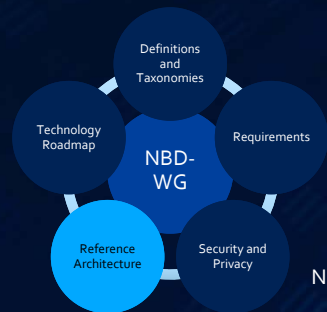
- Security and privacy are needed throughout lifecycle of processing stage and computing environment
- Collect various use cases and see if BDPP and BDML would work or not
- Identify what components are needed to enable BDPP and BDML



NBG-WG, Wo Chang, NIST, July 3, 2013



## Reference Architecture

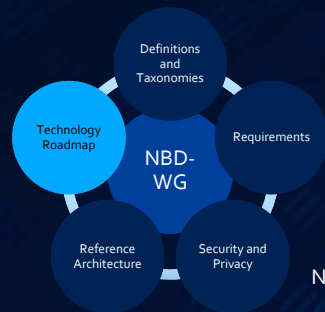


NBG-WG, Wo Chang, NIST, July 3, 2013

- Explore how other widely adopted standard architectures work, such as:
  - Hardware: Peripheral Component Interconnect (PCI)
  - Application: Browser Plug-ins
  - Protocol & Schema: HTTP, XML
  - Programming: Tcl / Tk (Tk can run many GUI platforms)
  - Markup: HTML5 – enables rich rendition on many systems / devices
- Identify what lessons can be adopted to enable BDPP and BDML



## Technology Roadmap



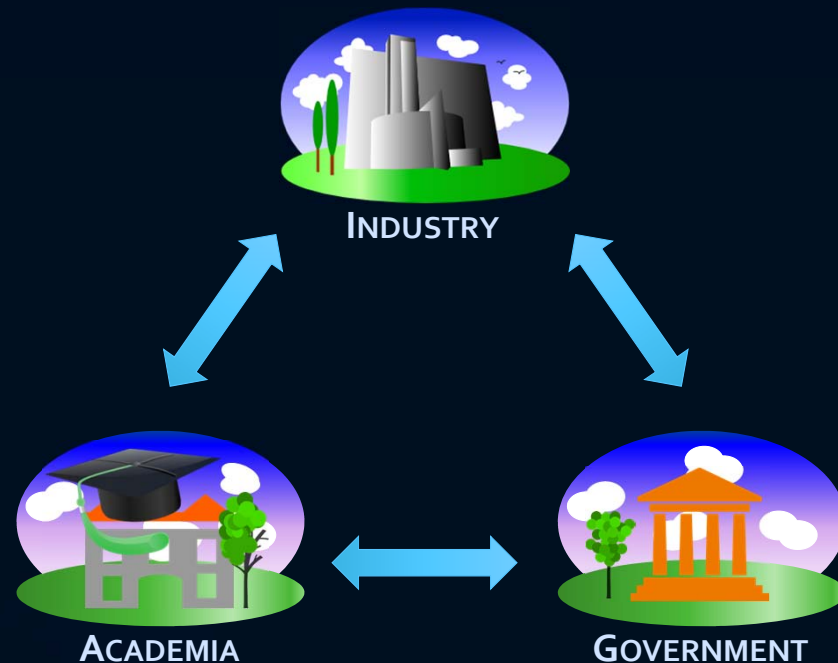
NBG-WG, Wo Chang, NIST, July 3, 2013

- Present forward vision for Big Data
- Present cohesive Big Data definitions, characteristics, and taxonomies
- Present Big Data roles and activities through use cases and scenarios
- Identify core Big Data requirements
- Present Reference Architecture to address those requirements
- Survey current Big Data standards and activities
- Perform gap analysis between vision and what's available
- Present recommendations
- Set standardization and adoption priorities for BDPP and BDML



# How can you help?

- Provide good use cases
- Provide feedback on Big Data direction
- Provide recommendations on Big Data adoption
- Others...



**Standards can benefit all users / systems!**

## DISCLAIMER:

This presentation is for suggestion only and is not a requirement for NBD-WG/Subgroups.