INFORMATION TECHNOLOGY LABORATORY

NIST Standard Enterprise Big Data Ecosystem

Wo Chang
Digital Data Advisor
ISO/IEC JTC1/WG 9 Big Data, Convenor
wchang@nist.gov

June 19, 2017

Agenda

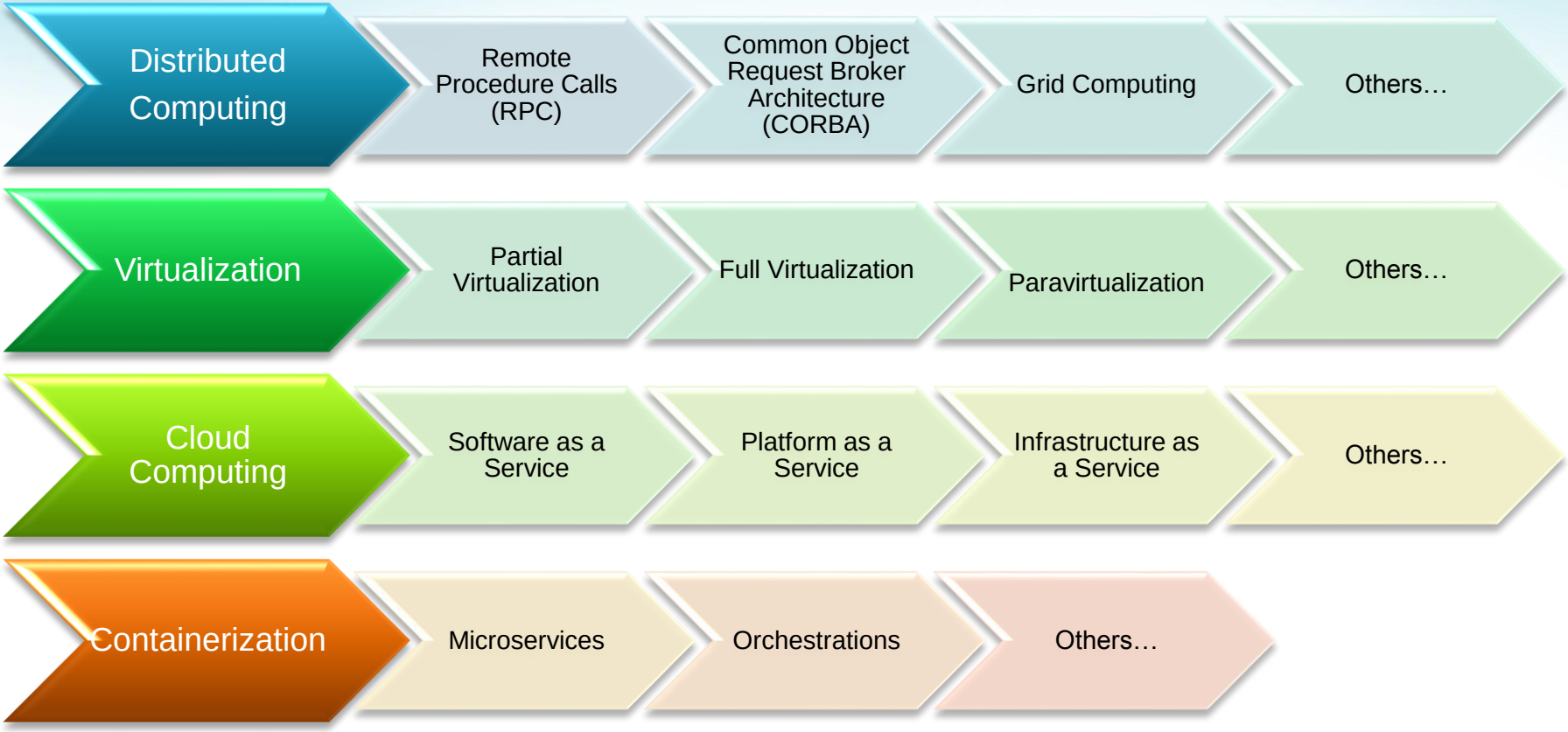
- Revisit Enterprise Computing
- What's the Computing Infrastructure Trend?
- What are the Big Data Architecture/Infrastructure Challenges?
- NIST Big Data Public Working Group Overview
- NIST Standard Enterprise Big Data Ecosystem
- NIST Big Data Interoperability Framework Version 1.0
- ISO/IEC JTC 1/WG 9 Big Data Standards Activities
- NIST Big Data Interoperability Framework Version 2.0
- NIST Big Data Reference Architecture (NBD-RA)
- Standard Big Data Analytics and Beyond

Revisit Enterprise Computing

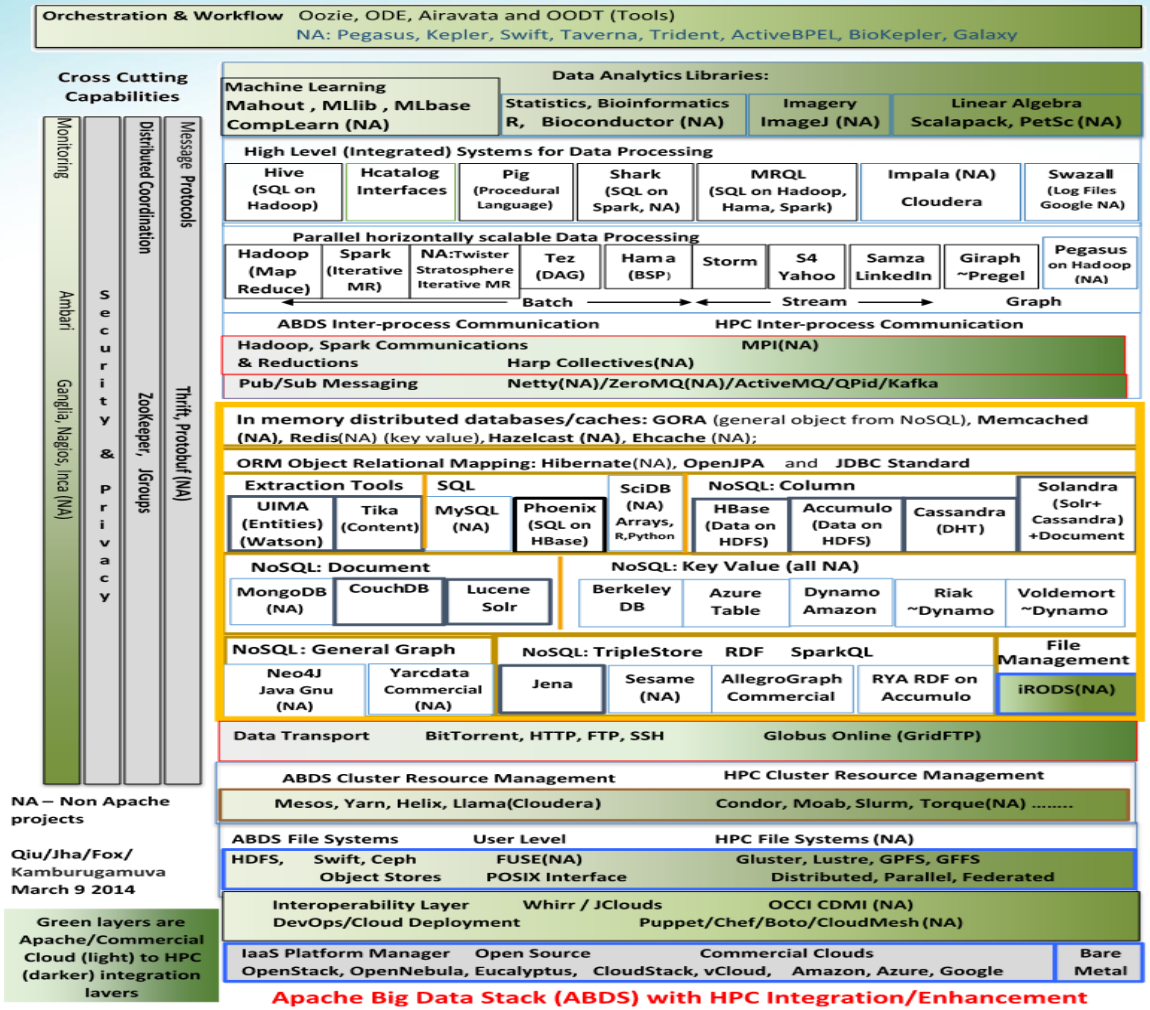
Enterprise computing is sometimes sold to business users as an entire platform that can be applied broadly across an organization and then further customized by users within each area. This means the *analytics, reporting, database management and other applications are standard across the system*, while the application packages being used and the data being accessed in each area will be different. In this sense, enterprise computing is a departure from finding single software solutions to specific business problems, such as inventory or accounting software. Instead, *enterprise computing is intended to offer integrated solutions to these problems*.

Source: https://www.techopedia.com/definition/27854/enterprise-computing

What's the Computing Infrastructure Trend?



Big Data Architecture/Infrastructure – Challenges (Computing Stack)

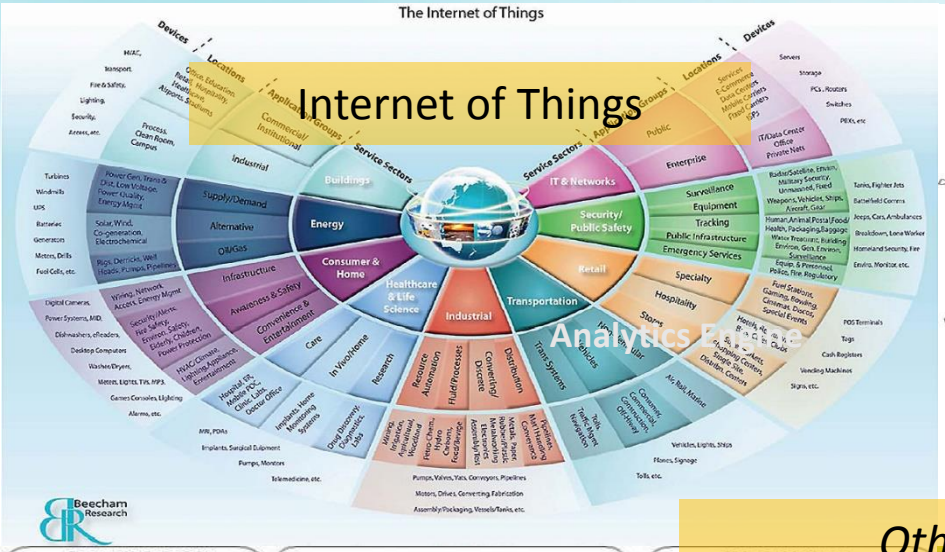


NIST Standard Enterprise Big Data Ecosystem, Wo Chang, NIST/ITL, June 19, 2017

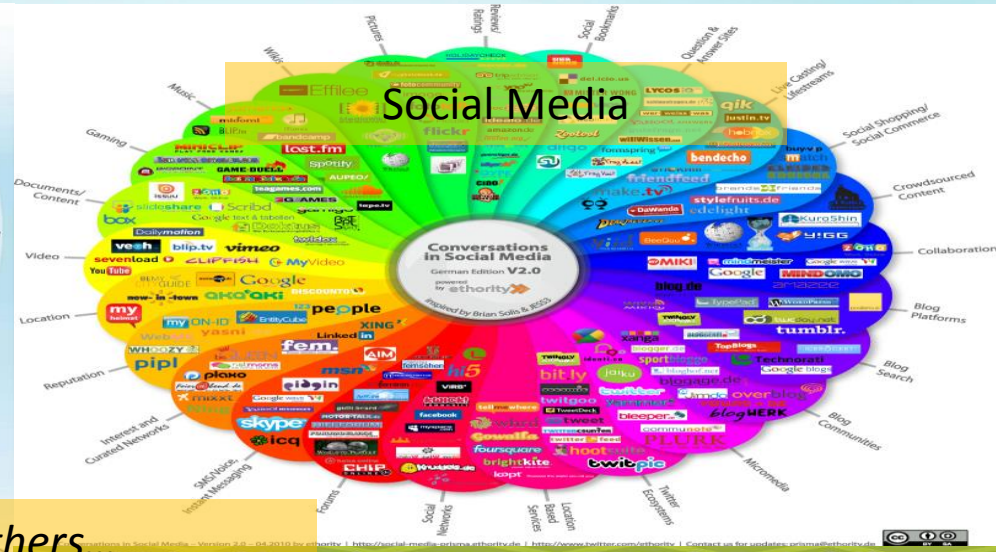
Big Data Architecture/Infrastructure – Challenges (Analytics Stack)

Source: http://1.bp.blogspot.com/-PKiTQa0mrn4/T_mGb6AI3yI/AAAAAAAAA8Q/TtH7xyjQ3FA/s640/analytics+tools+landscape.bmp

Big Data Architecture/Infrastructure – Challenges (Applications Stack)



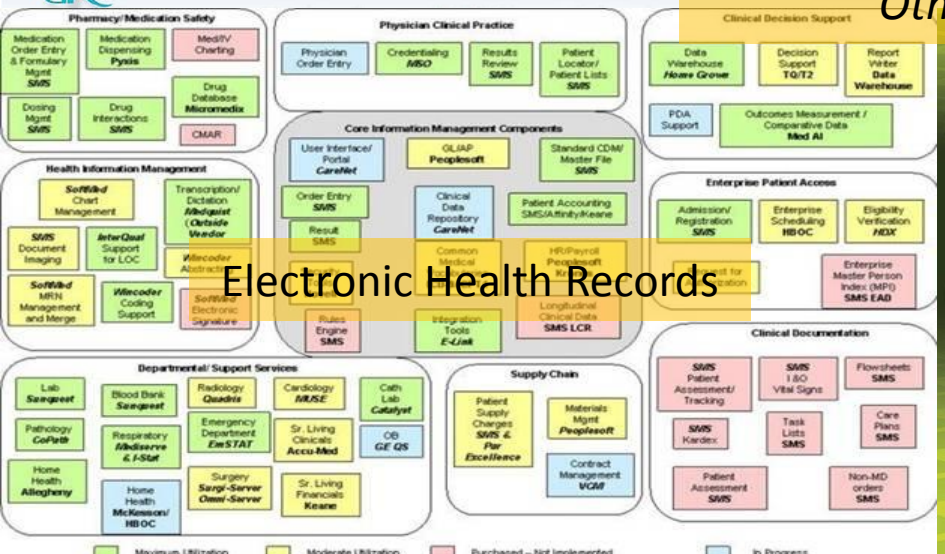
Internet of Things



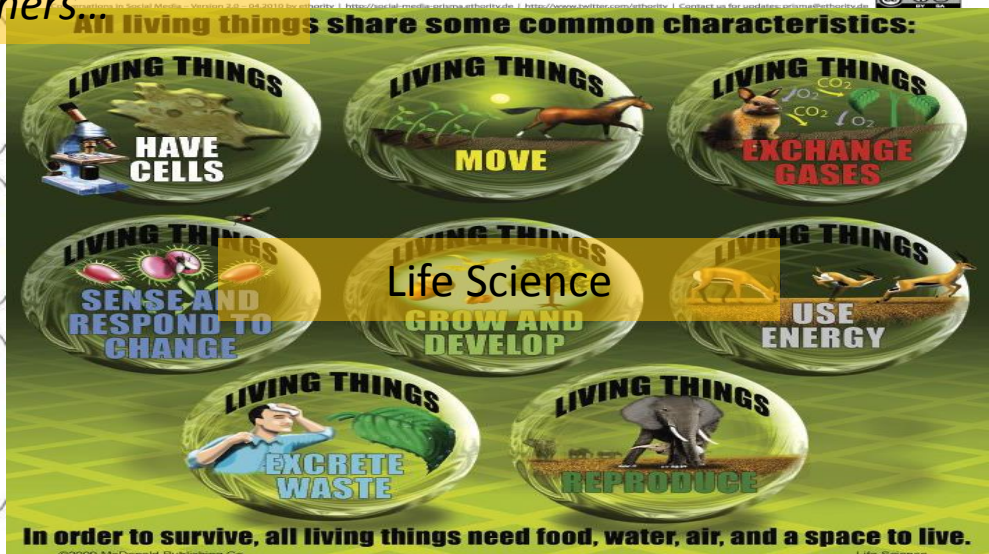
Social Media

Analytics Engine

Others…

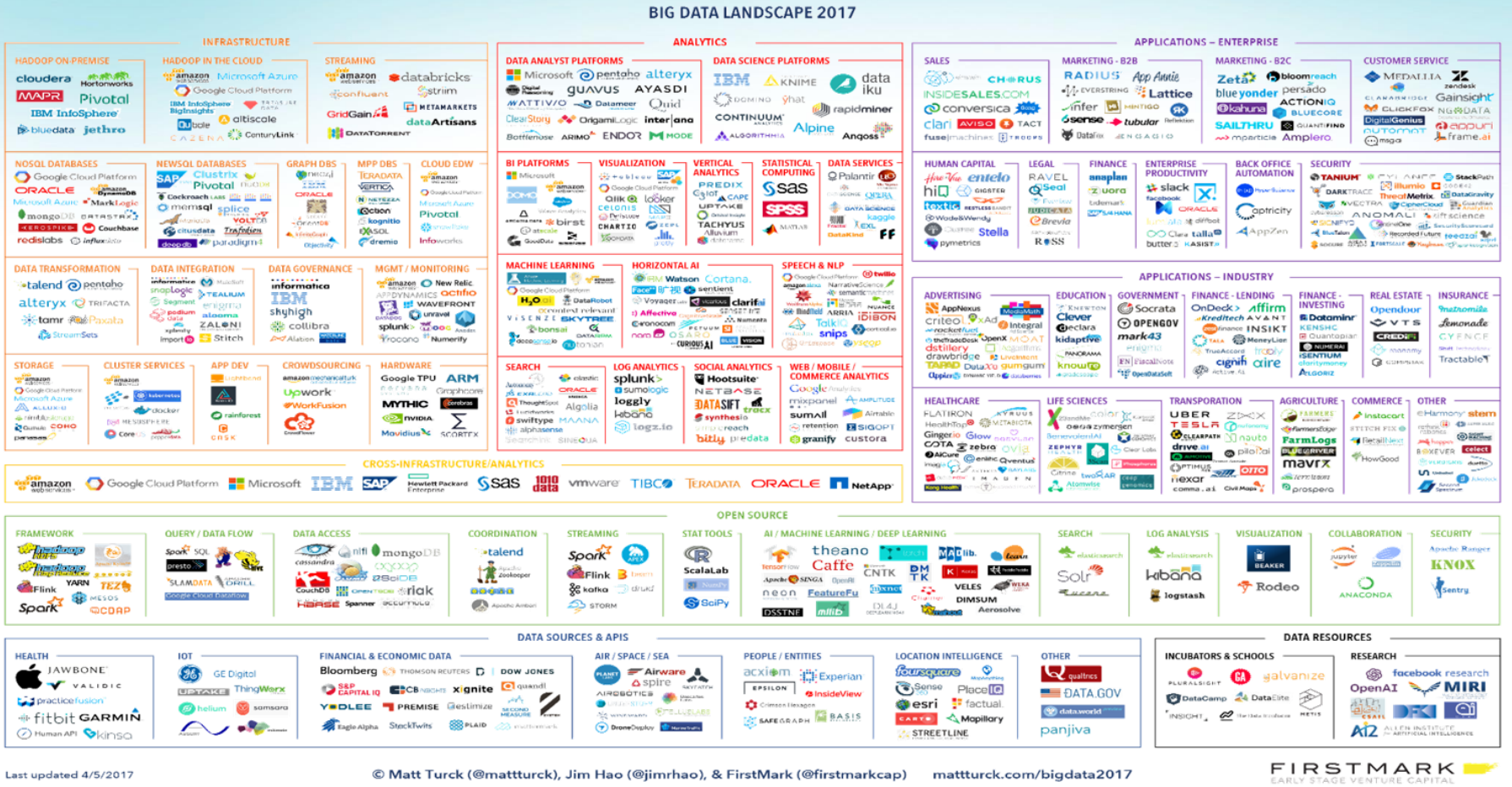


Electronic Health Records



Life Science

Big Data Architecture/Infrastructure – Challenges (Integration)



Source: http://mattturck.com/wp-content/uploads/2017/04/Big-Data-Landscape-2017-Matt-Turck-FirstMark.png

NIST Big Data Public Working Group Overview

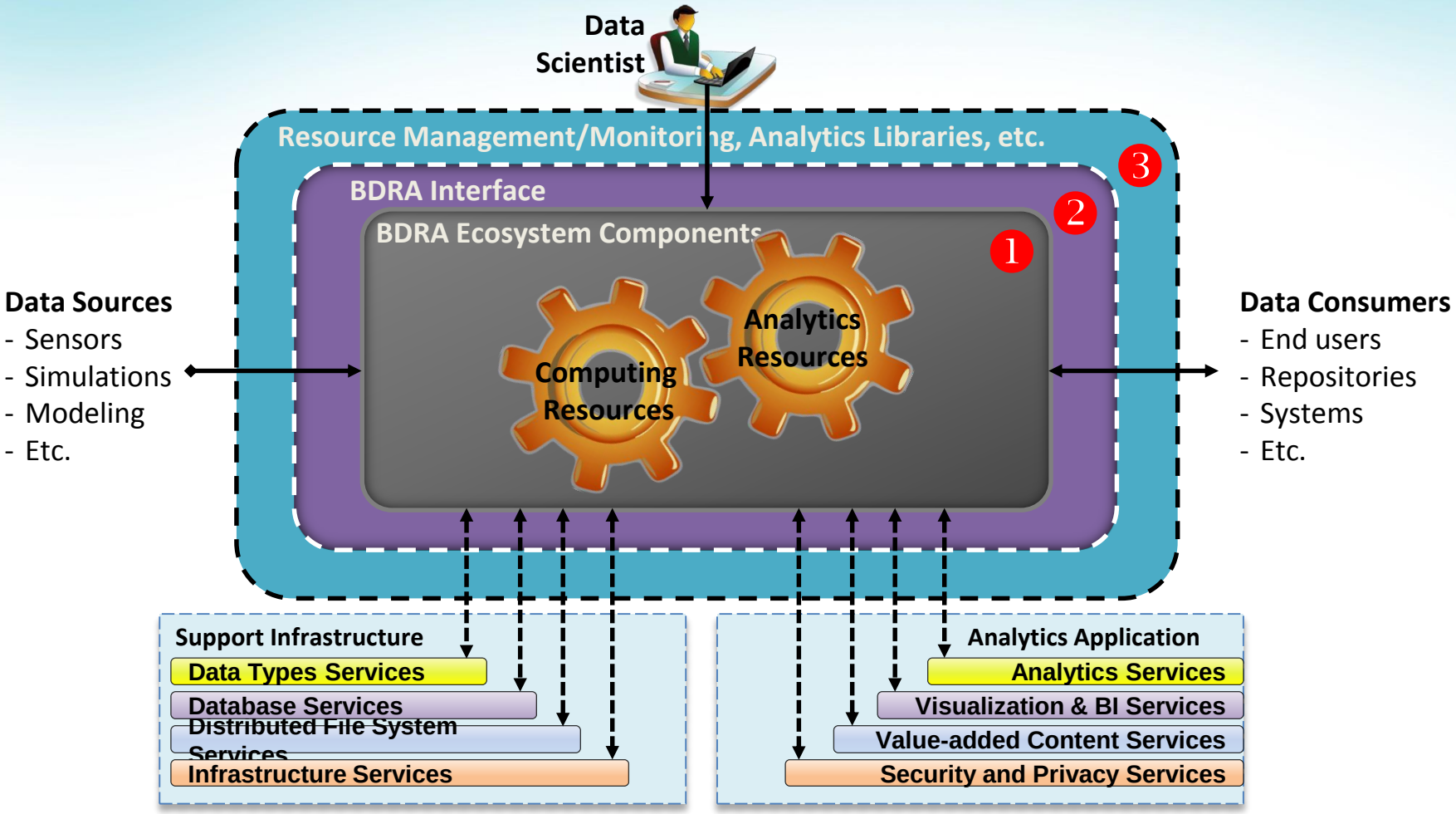
Goal

Work with industry, academia and government to create a consensus-based extensible NIST Big Data Interoperability Framework (NBDIF) which is a vendor-neutral, technology- and infrastructure-independent ecosystem. It can enable Big Data stakeholders (e.g. data scientists, researchers, etc.) to utilize the best available analytics tools to process and derive knowledge through the use of standard interfaces between swappable architectural components.

Approach – Three Stages (refers to three versions)

1. Identify the high-level NIST Big Data Reference Architecture (NBDRA) key components, which are technology-, infrastructure-, and vendor-agnostic [Done]
2. Define general interfaces between the NBDRA key components [Ongoing]
3. Validate the NBDRA by building Big Data general applications through the general interfaces.

NIST Standard Enterprise Big Data Ecosystem



NIST Standard Enterprise Big Data Ecosystem, Wo Chang, NIST/ITL, June 19, 2017

NIST Big Data Interoperability Framework Version 1

Deliverable: Stage-1 – High-level Reference Architecture

https://bigdatawg.nist.gov/V1_output_docs.php (Sep. 16, 2015)

- NIST SP1500-1: Definitions
- NIST SP1500-2: Taxonomies
- NIST SP1500-3: Use Cases & Requirements
- NIST SP1500-4: Security & Privacy
- NIST SP1500-5: Architecture Survey – White Paper
- NIST SP1500-6: Reference Architecture
- NIST SP1500-7: Standards Roadmap

ISO/IEC JTC 1/WG 9 Big Data Standards Activities

ISO/IEC JTC 1/WG 9 Working Group on Big Data (Jan. 2015 – now)

- 180+ from 26 NBs: Australia, Austria, Brazil, Canada, China, Finland, France, Germany, India, Ireland, Israel, Japan, Korea, Luxembourg, Mexico, Netherlands, Norway, Russian Federation, Saudi Arabia, Singapore, Slovenia, South Africa, Spain, Sweden, UK, US
- Current Projects
  - **ISO/IEC 20546 Information technology – Big data – Definition and vocabulary** (Committee Draft #3 as June 2017)
  - **ISO/IEC 20547 Information Technology – Big data Reference architecture (5 Parts as June 2017)**
    - Part 1: (TR) Framework and Application Process (2nd WD)
    - Part 2: (TR) Use Cases and Derived Requirements (Recommended for Publication)
    - Part 3: (IS) Reference Architecture (5th WD)
    - Part 4: (IS) Security and Privacy Fabric (2nd ED, under SC 27/WG 4)
    - Part 5: (TR) Standards Roadmap (Recommendation for Publication)
- ISO/IEC Liaisons: SC 6/WG 7, SC 27, SC 29, SC 32, SC 36, SC 38, SC 39, ISO/TC 69, ISO/TC 204, ITU-T SG13, IIC, OGC, BDVA (coming up)

ISO/IEC JTC 1/WG 9 Big Data Standards Activities

1st International Workshop on Big Data Standards, March 7, 2016, Dublin, Ireland

16 Speakers, One Panel Discussion (4 panelists), 80+ Registered, 50+ attended

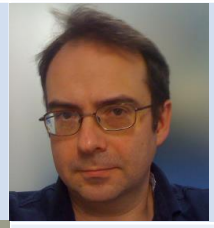
Leo Clancy
IDA Head of Technology
Welcome, Speakers, Delegates and Visitors

Ian Cowan
NSAI Ireland
ICT and Big Data Standardization in Ireland

Wo Chang
ISO IEC JTC1 WG9 Convenor
WG9 Big Data Standard

Jean Stride
BSI UK
Big Data: A Market Strategy

Maurice Buckley
CEO NSAI & Vice-President Technical CENELEC
Supporting the Digital Single Market through Standardization: A European and Irish perspective on Industry 4.0 and Big Data



Dave Lewis
Spokes Director - ADAPT Ireland
Standards for the Multilingual Web

Colette Maloney
HoU Smart Cities and Sustainability
DG CONNECT
EU Smart Cities and Big Data

Edward Curry
VP Big Data Value Assoc
Big Data Value PPP: A Standardisation Opportunity for Europe

ISO/IEC JTC 1/WG 9 Big Data Standards Activities

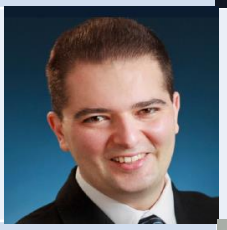
1st International Workshop on Big Data Standards, March 7, 2016, Dublin, Ireland

16 Speakers, One Panel Discussion (4 panelists), 80+ Registered, 50+ attended

Phil Archer
W3C UK
Big Data Europe, an infrastructure for all

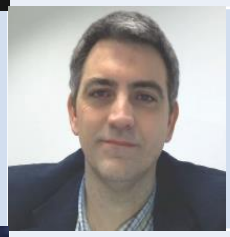
Rob Brennan
EU ALIGNED
ALIGNED - Bringing together Software and Data Engineering for Data-intensive Systems

Wael William Diab
FutureWei
Big Data Ecosystem



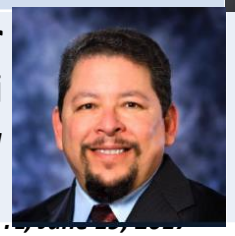
Dudley Dolan
Q-VALIDUS
Towards a CEN Workshop on Big Data

Moez Draief
Huawei France
Parsimonious network monitoring



Ismael Caballero
UCLM Spain
Data Quality for Big Data

John Strassner
Futurewei
Engineering Value from Big Data

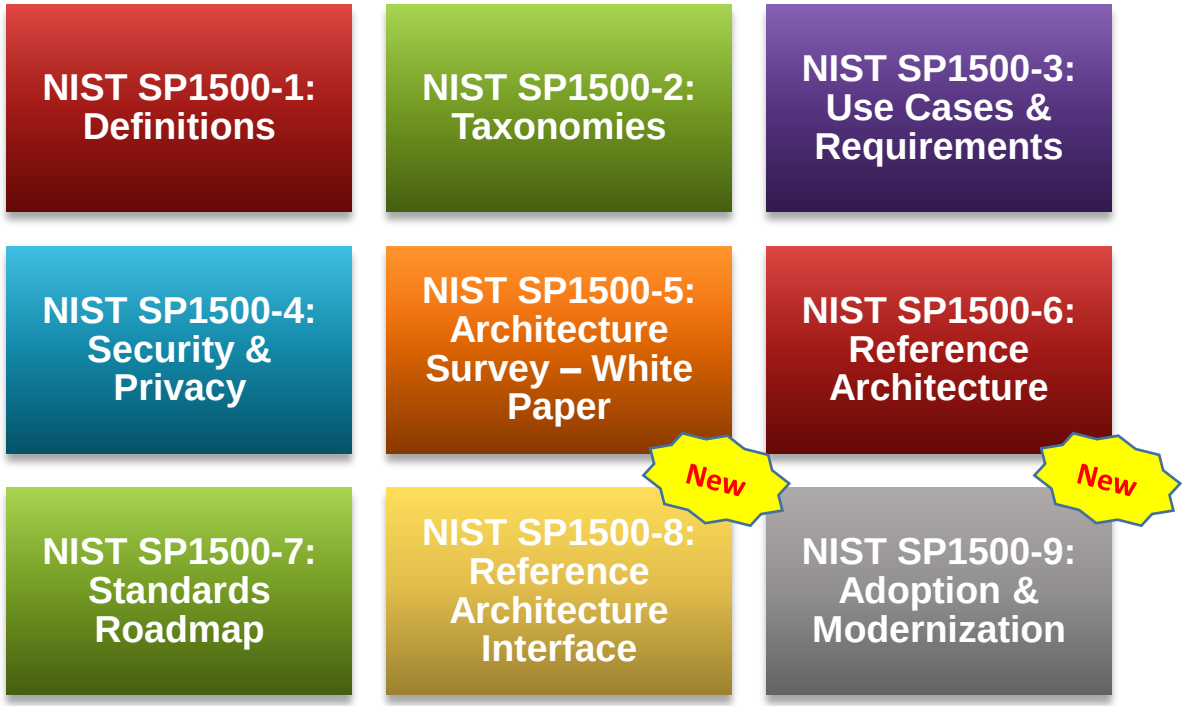


Souleiman Hasan
Insight Ireland
Big Data Technical Priorities

NIST Big Data Interoperability Framework Version 2

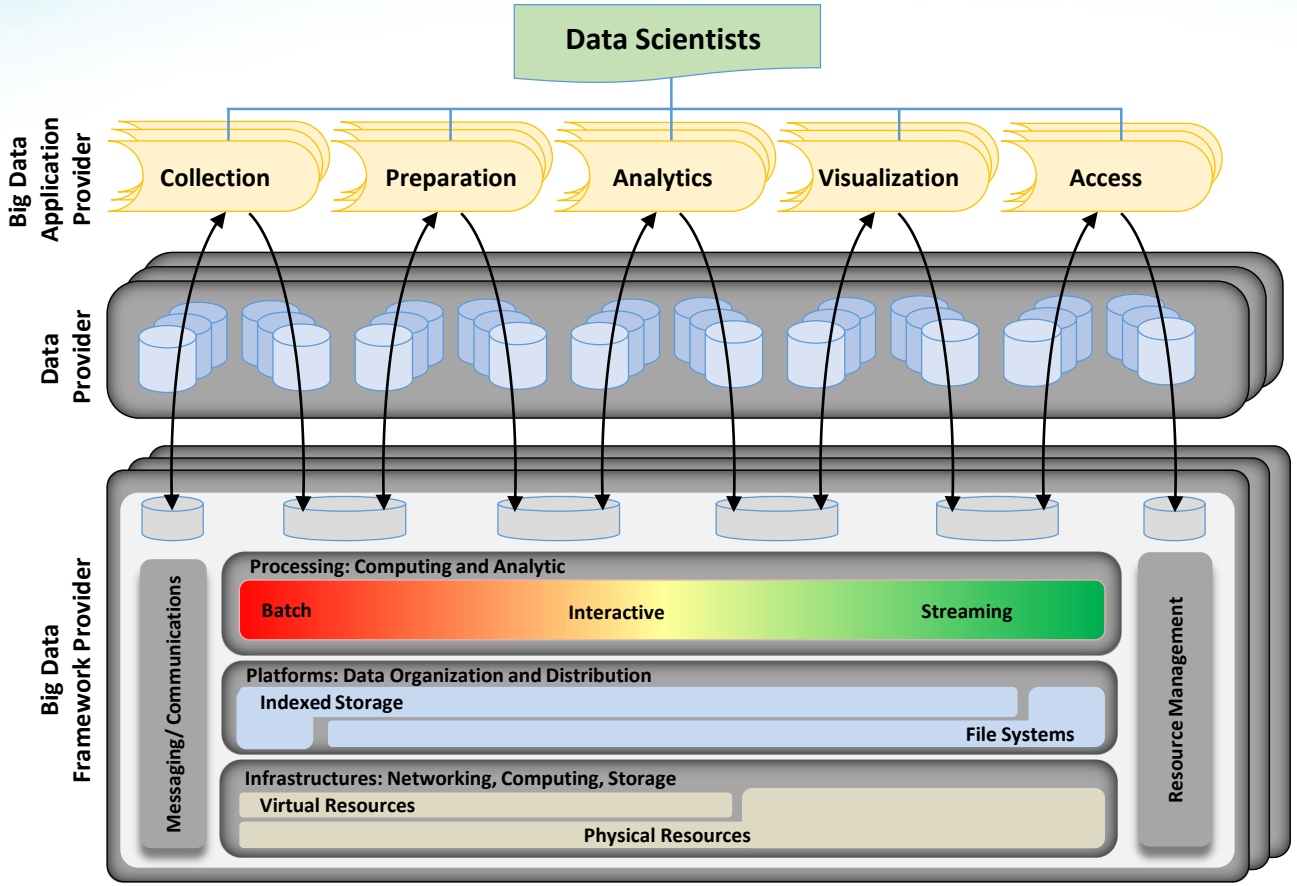
Deliverable: Stage-2 – Reference Architecture Interface

https://bigdatawg.nist.gov/V2_output_docs.php (Nov. 2017)



NIST Big Data Reference Architecture (NBD-RA)

Enable data scientists, engineers, researchers, etc. to increase productive and enhance quality in data science through standard modularized Big Data Analytics tools.



NIST Standard Enterprise Big Data Ecosystem, Wo Chang, NIST/ITL, June 19, 2017

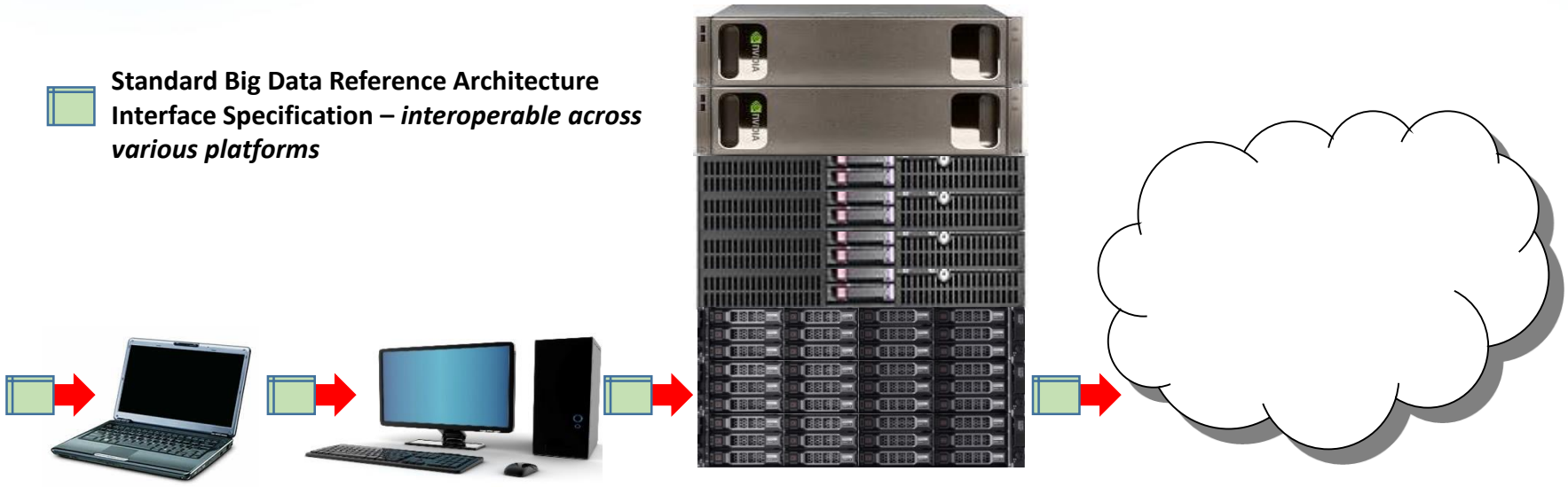
Standard Big Data Analytics and Beyond

Enable Big Data analytics tools for *interoperability, portability, reusability, and extensibility*.

Practical Aspect: Analytics tools can be *reusable, deployable, and operational* (max. use of resources)

[similar approach as to HTML, PCI bus, etc.]

Standard Big Data Reference Architecture Interface Specification – *interoperable across various platforms*



Laptop | **Desktop/Server** | **……** **Data Center** **Many CPUs/Cores/GPUs/ FPGAs/ASICs/accelerators** | **Cloud**