NIST Big Data Reference Architecture for Analytics and Beyond

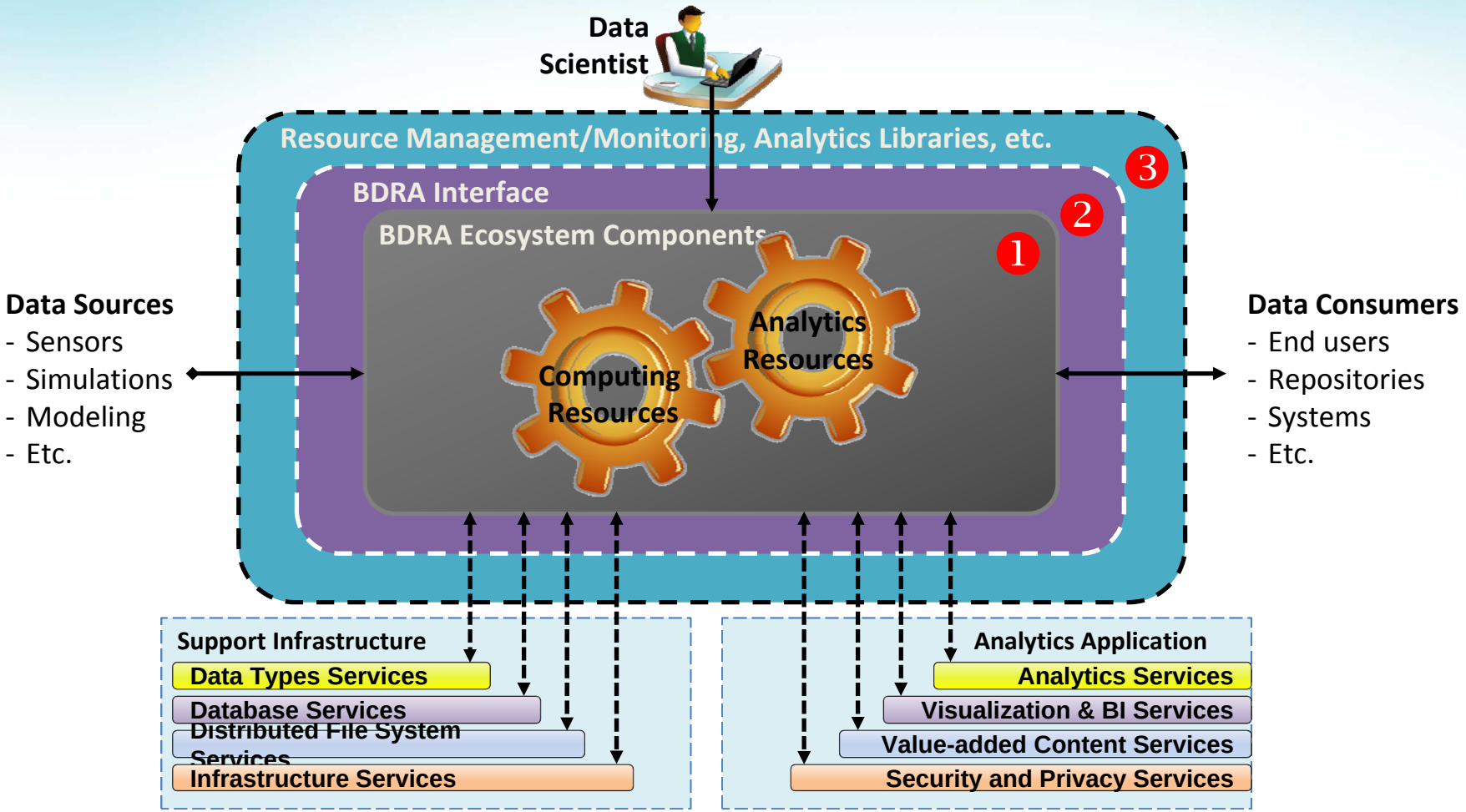
Wo Chang
Digital Data Advisor
wchang@nist.gov

June 2, 2017

Agenda

- Preposition: Assume NBD-RA is Established, then What?
- Revisit Enterprise Computing
- What's the Computing Infrastructure Trend?
- What are the Big Data Architecture/Infrastructure Challenges?
- Goals for Big Data Analytics and Beyond
- Complement to IEEE Big Data Initiative & Direction

Preposition: Assume NBD-RA Established, then What?



Revisit Enterprise Computing

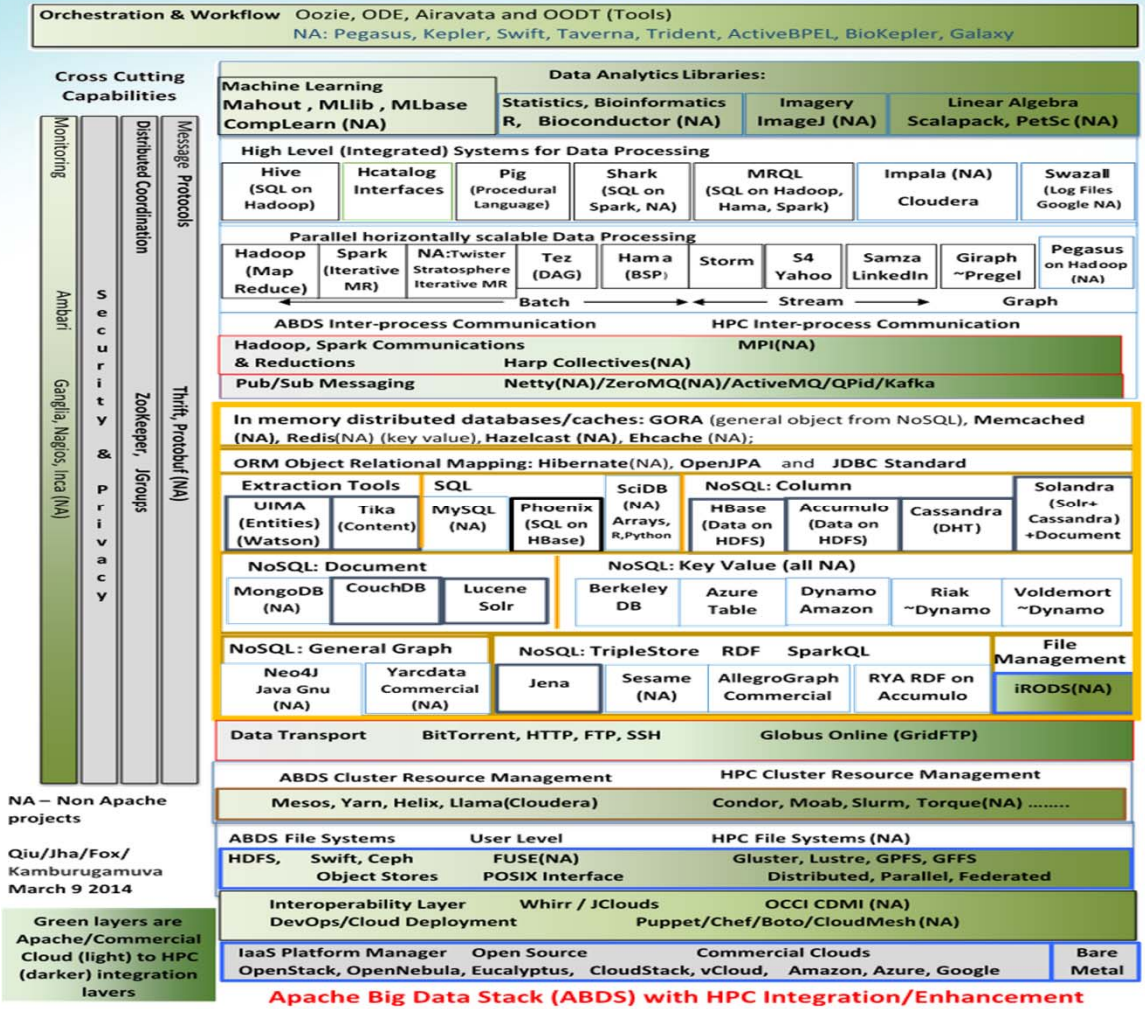
Enterprise computing is sometimes sold to business users as an entire platform that can be applied broadly across an organization and then further customized by users within each area. This means the *analytics, reporting, database management and other applications are standard across the system*, while 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?

Distributed Computing	Remote Procedure Calls (RPC)	Common Object Request Broker Architecture (CORBA)	Grid Computing	Others…
Virtualization	Partial Virtualization	Full Virtualization	Paravirtualization	Others…
Cloud Computing	Software as a Service	Platform as a Service	Infrastructure as a Service	Others…
Containerization	Microservices	Orchestrations	Others…	

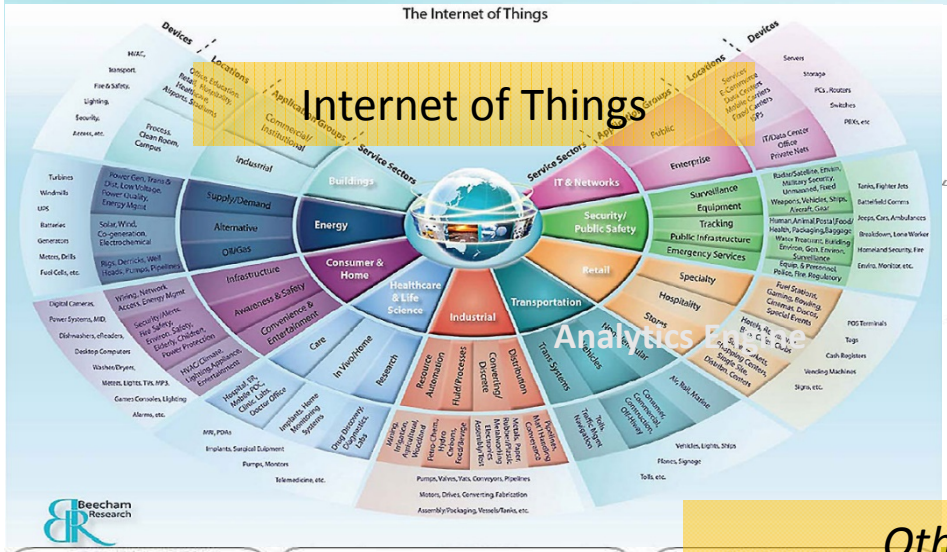
Big Data Architecture/Infrastructure – Challenges (Computing Stack)



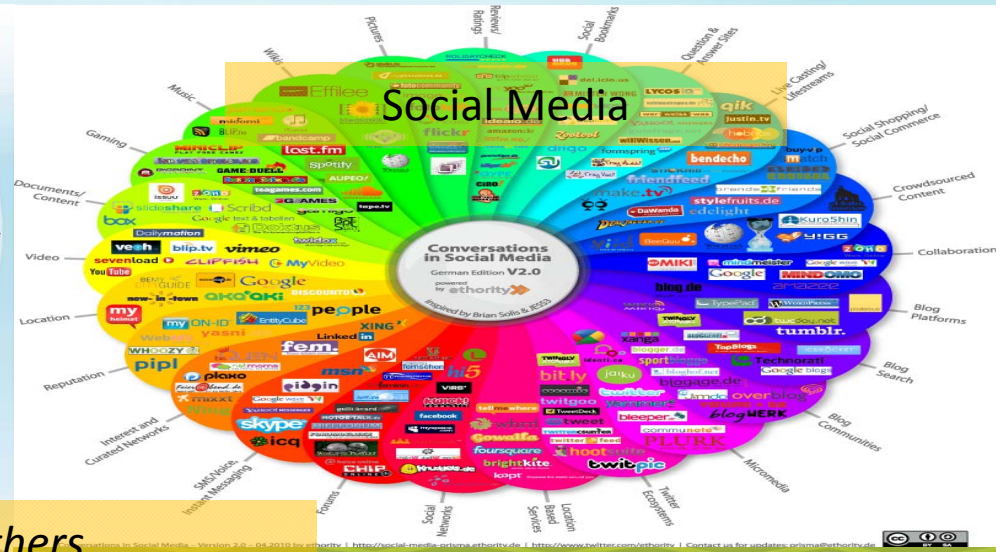
Big Data Architecture/Infrastructure – Challenges (Analytics Stack)

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

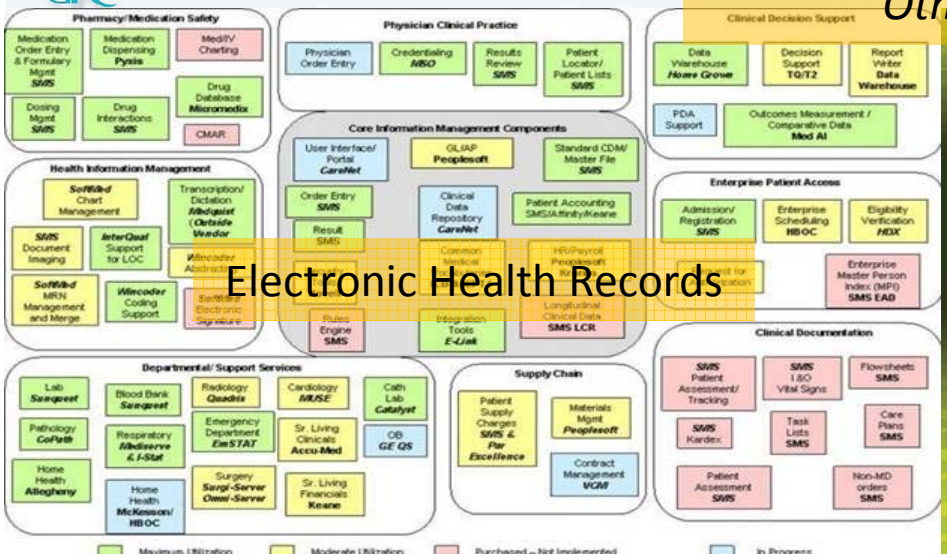
Big Data Architecture/Infrastructure – Challenges (Applications Stack)



Internet of Things

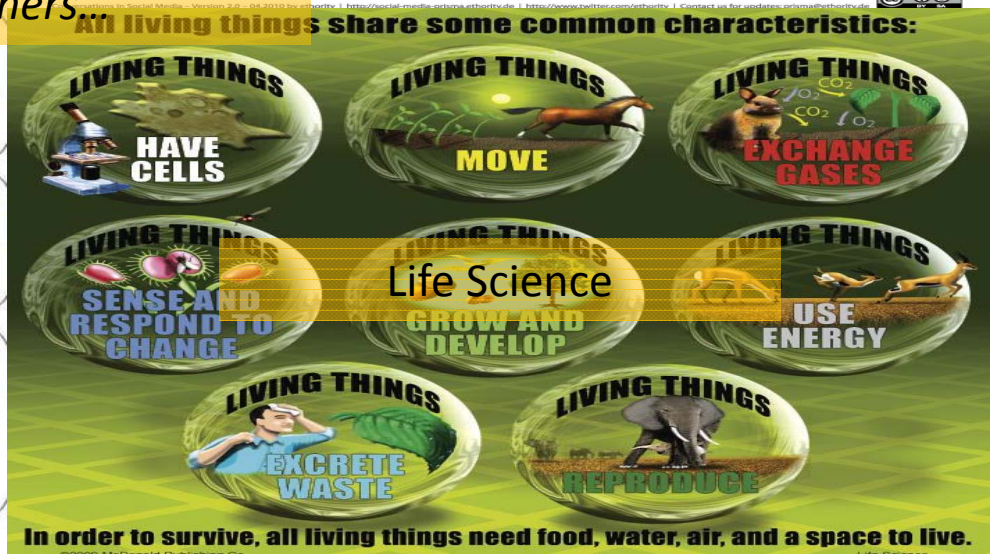


Social Media



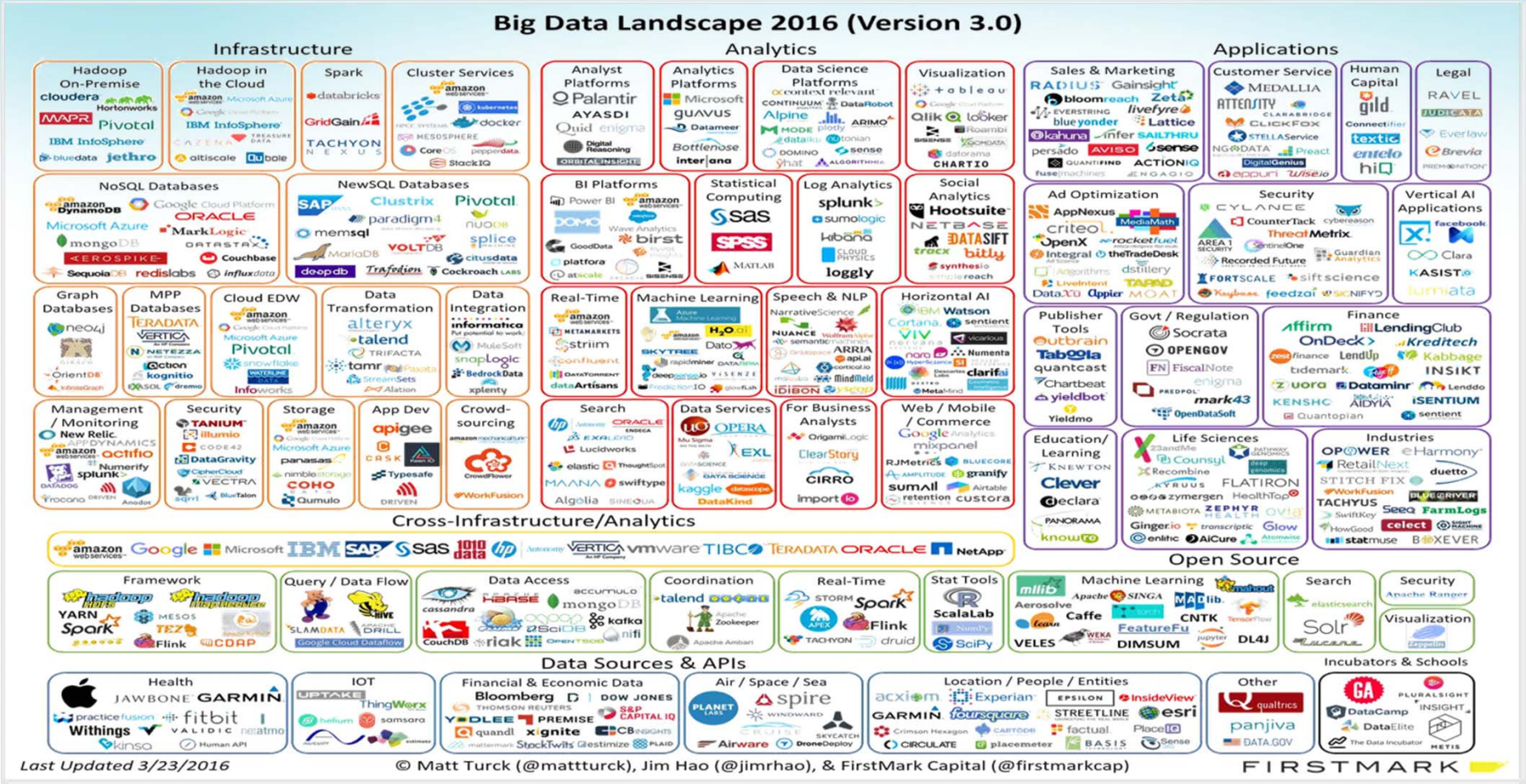
Electronic Health Records

Others…



Life Science

Big Data Architecture/Infrastructure – Challenges (Integration)



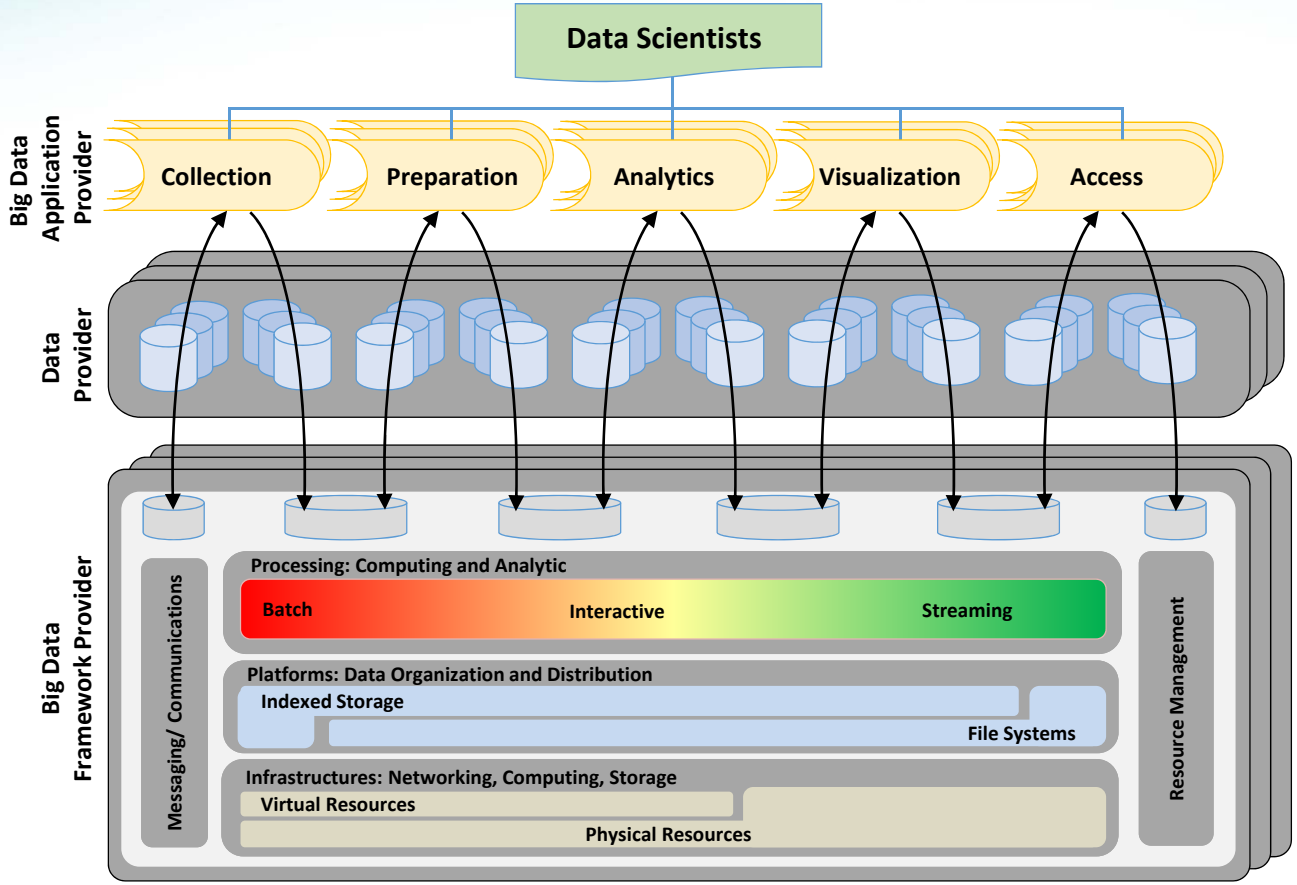
Source: https://www.interana.com/wp-content/uploads/2017/03/Big-Data-Landscape-2016-v18-FINAL.jpg

Goals for 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.]

Complement to IEEE Big Data Initiative & Direction

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