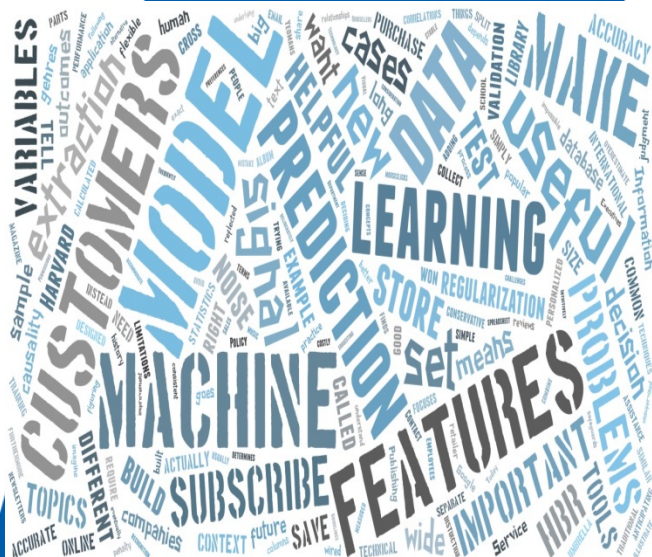


THE VALUE OF PERFORMANCE.
NORTHROP GRUMMAN

Chief Medical Officer

IOM Regulatory Science Workshop

October 21, 2015



Key Need – Ability to Address the 5 V's of Big Data Analytics

Big Data Analytics: The curation and examination of massive amounts of structured and unstructured data to elucidate unknown or unproven patterns, correlations, and trends.

The 5 V's:

Volume

How much?

90% of the world's data has been created in the last 2 years

Velocity

How fast?

More than 170 websites are created every minute

Variety

How many?

70% of the digital universe (900 exabytes) is generated by different users

Veracity

How true?

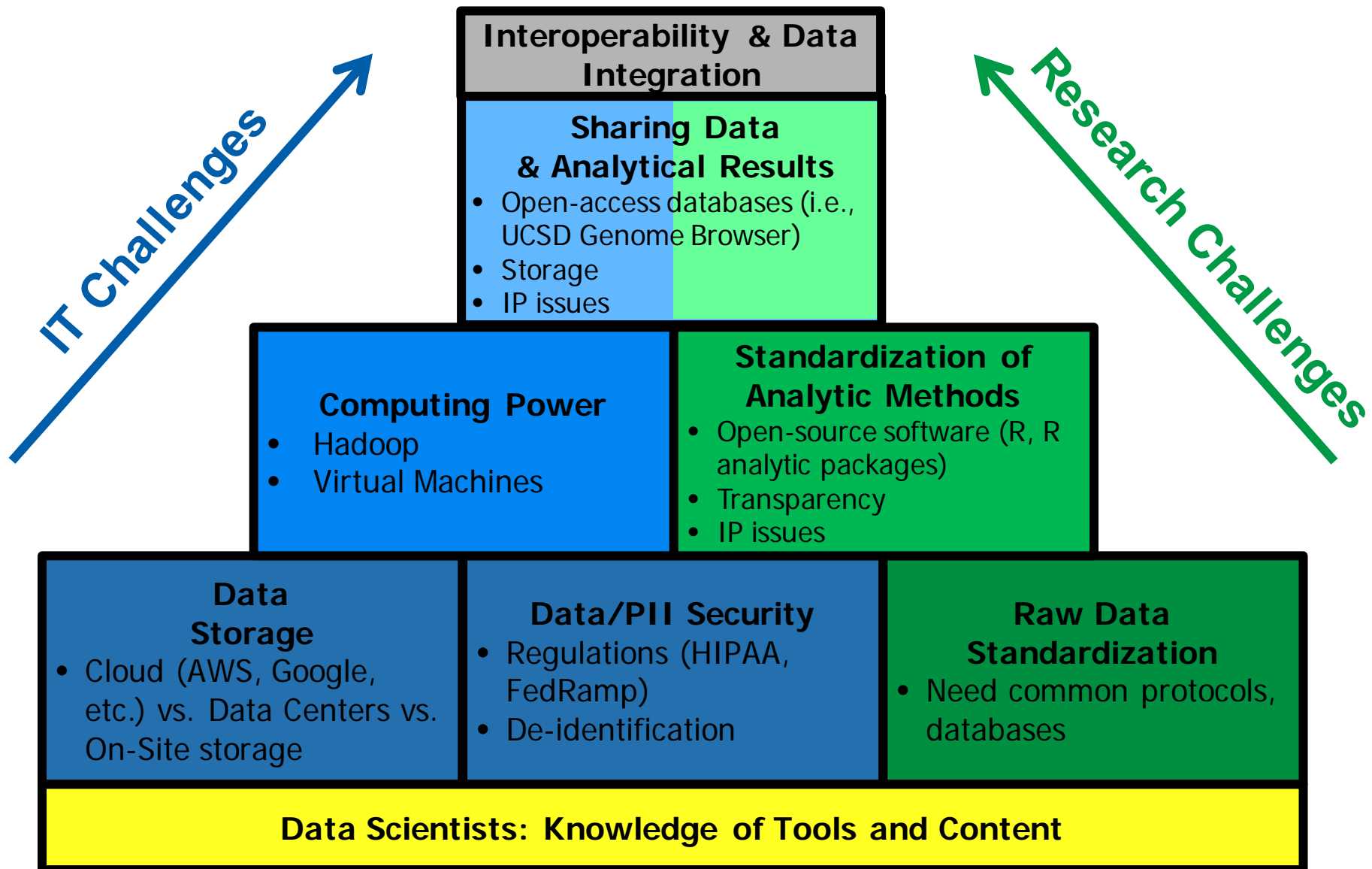
1 in 3 business leaders don't trust the information they use to make decisions

Value

How useful?

65% of senior execs say management decisions are increasingly based on hard analytic information

Data Scientists are the Foundation for Solving Today's Big Data Challenges




Knowledge, Skills, and Abilities of a Data Scientist

MODERN DATA SCIENTIST

Data Scientist, the sexiest job of the 21st century, requires a mixture of multidisciplinary skills ranging from an intersection of mathematics, statistics, computer science, communication and business. Finding a data scientist is hard. Finding people who understand who a data scientist is, is equally hard. So here is a little cheat sheet on who the modern data scientist really is.

MATH & STATISTICS

- ☆ Machine learning
- ☆ Statistical modeling
- ☆ Experiment design
- ☆ Bayesian inference
- ☆ Supervised learning: decision trees, random forests, logistic regression
- ☆ Unsupervised learning: clustering, dimensionality reduction
- ☆ Optimization: gradient descent and variants



PROGRAMMING & DATABASE

- ☆ Computer science fundamentals
- ☆ Scripting language e.g. Python
- ☆ Statistical computing packages, e.g., R
- ☆ Databases: SQL and NoSQL
- ☆ Relational algebra
- ☆ Parallel databases and parallel query processing
- ☆ MapReduce concepts
- ☆ Hadoop and Hive/Pig
- ☆ Custom reducers
- ☆ Experience with xaaS like AWS

DOMAIN KNOWLEDGE & SOFT SKILLS

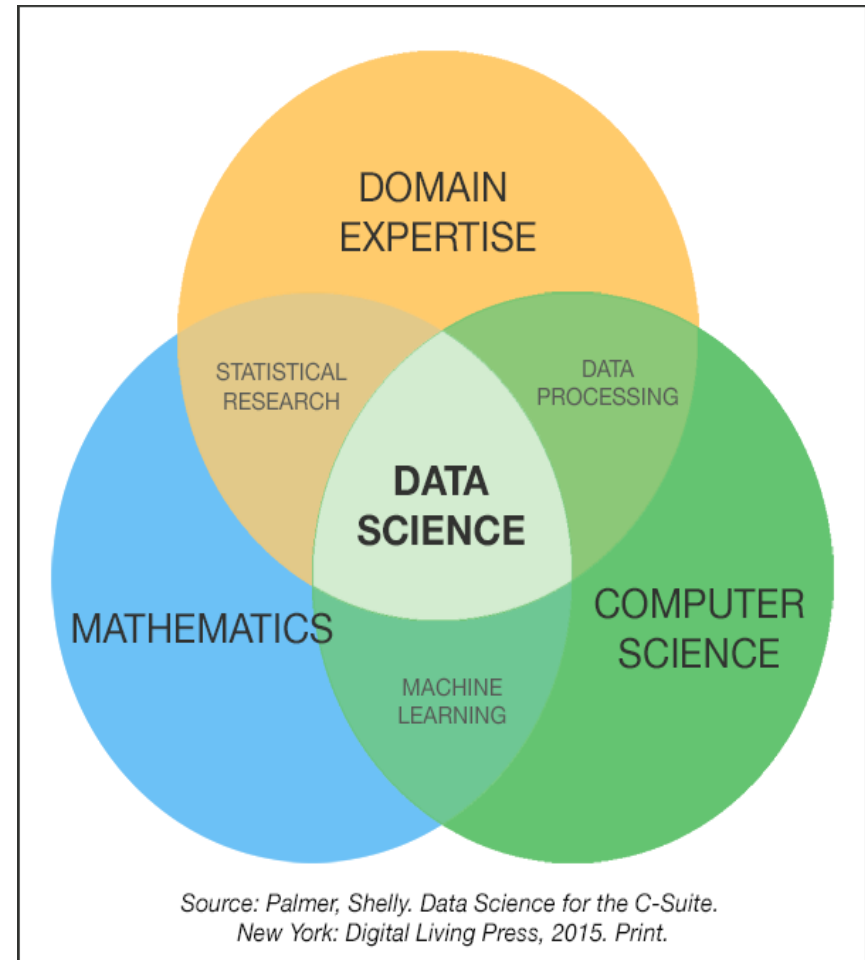
- ☆ Passionate about the business
- ☆ Curious about data
- ☆ Influence without authority
- ☆ Hacker mindset
- ☆ Problem solver
- ☆ Strategic, proactive, creative, innovative and collaborative

COMMUNICATION & VISUALIZATION

- ☆ Able to engage with senior management
- ☆ Story telling skills
- ☆ Translate data-driven insights into decisions and actions
- ☆ Visual art design
- ☆ R packages like ggplot or lattice
- ☆ Knowledge of any of visualization tools e.g. Flare, D3.js, Tableau

MarketingDistillery.com is a group of practitioners in the area of e-commerce marketing. Our fields of expertise include: marketing strategy and optimization; customer tracking and on-site analytics; predictive analytics and econometrics; data warehousing and big data systems; marketing channel insights in Paid Search, SEO, Social, CRM and brand.

**Marketing
DISTILLERY**
(c) Krzysztof Zawadzki



Data Scientists highly valued - must be trained and proficient in multiple inter-related disciplines

Needs and Challenges

Develop a data science talent pool that can capitalize on the growing data ecosystem:

- Unlock the hidden value by making more and better use of big data
- Optimize performance and quality through more accurate and detailed data collection and analysis
- Segment patients and consumers more precisely
- Improve decision-making through sophisticated analysis
- Mature the industry along the sophistication continuum from descriptive, to diagnostic, to predictive, to prescriptive analytics

Training Options

Traditional academic graduate degree and certificate programs

- Programs in Data Science, Analytics, Data Mining, Predictive Analytics
- Interdisciplinary institutes and centers of excellence, such as Michigan Institute for Data Science (MIDAS)

Non-traditional online learning

- Coursera, EdX, MIT OpenCourseware

Industry training

- NIH Data Science Workforce Development Center:
 - Development of open course materials/online courses
 - Hosting of in-person courses
 - Discovery of educational resources

Corporate training

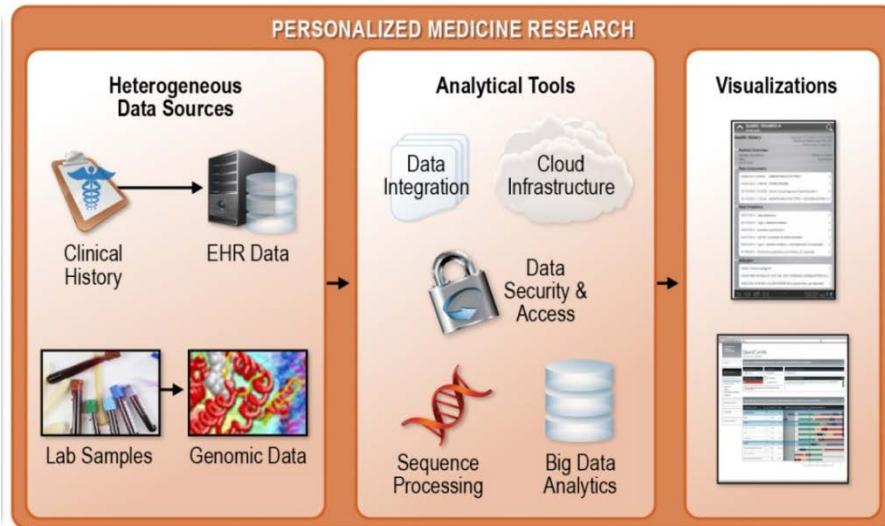
- MIT Sloan study: 63% of companies are turning to on-the-job training

US will have shortage of 140-190K people with deep analytical skills

(McKinsey Global Institute)

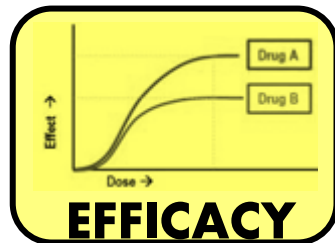
Impact of Data Scientists and Big Data Analytics

Data Scientists Drive Big Data Solutions



Potential Value/Impact for Regulatory Science:

- Reveal unknown drug interactions
- Link genetics with treatment and/or response
- Pharmaceutical production—improve product safety, reduce production costs
- Combat healthcare fraud
- Track and identify adverse drug events



Key Personnel in harnessing multiple sources of data to help achieve health and scientific goals

THE VALUE OF PERFORMANCE.

NORTHROP GRUMMAN

A blue curved line graphic that starts below the company name and sweeps upwards and to the right, ending under the right side of the name.