



State of the Union The World of Data

Northern California Oracle Users Group – Spring 2020

John Kanagaraj, Sr. MTS, Enterprise Data Platform • 21-May-2020

Agenda

Introduction

“Data is the New Oil”

Challenges & Influencing Factors

Responding to the Challenges

Wrap up: Q & A

Setting the Stage

- Acknowledging the Current Situation : A Black Swan event
 - Worldwide Pandemic: Health and Economies affected
 - Balancing collective health vs economy
- Self Survey
 - Challenges in Operations and Development
 - Proliferation of Technologies
 - Adapting, Learning and Growing
- Scope
 - Trends and Threats
 - Opportunities and Growth
- What we will not get into
 - Predicting the future!



https://en.wikipedia.org/wiki/Black_swan_theory

About PayPal



**300M+ Accounts
20M+ Merchants**



**3B+ Payment
Transactions
40 Payment transactions
/account trailing 12
months**



800TB+ Database

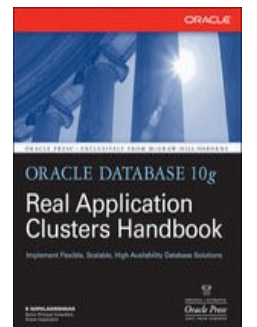
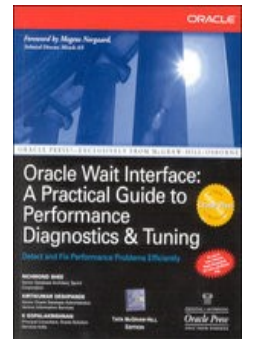
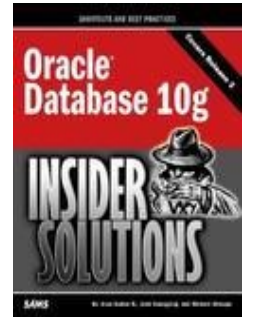


**1.5M SQL Executions/Sec
on a Database Group**

- 300+M Active Accounts with 20M+ Merchants
- 3B Payment transactions
- 40 Payment transactions per active account in last 12 months
- Continuous Double- Digit Growth
- Hundreds of Databases with Hundreds of PBs of total storage
- Biggest DB - 800+ TB
- 1.5M SQL Executions/Sec on a DB Group
- 4TB Redo per hour on a Busy DB Group
- >20k Inserts/sec ,>50k selects/sec
- Robust Big Data and AI/ML platforms

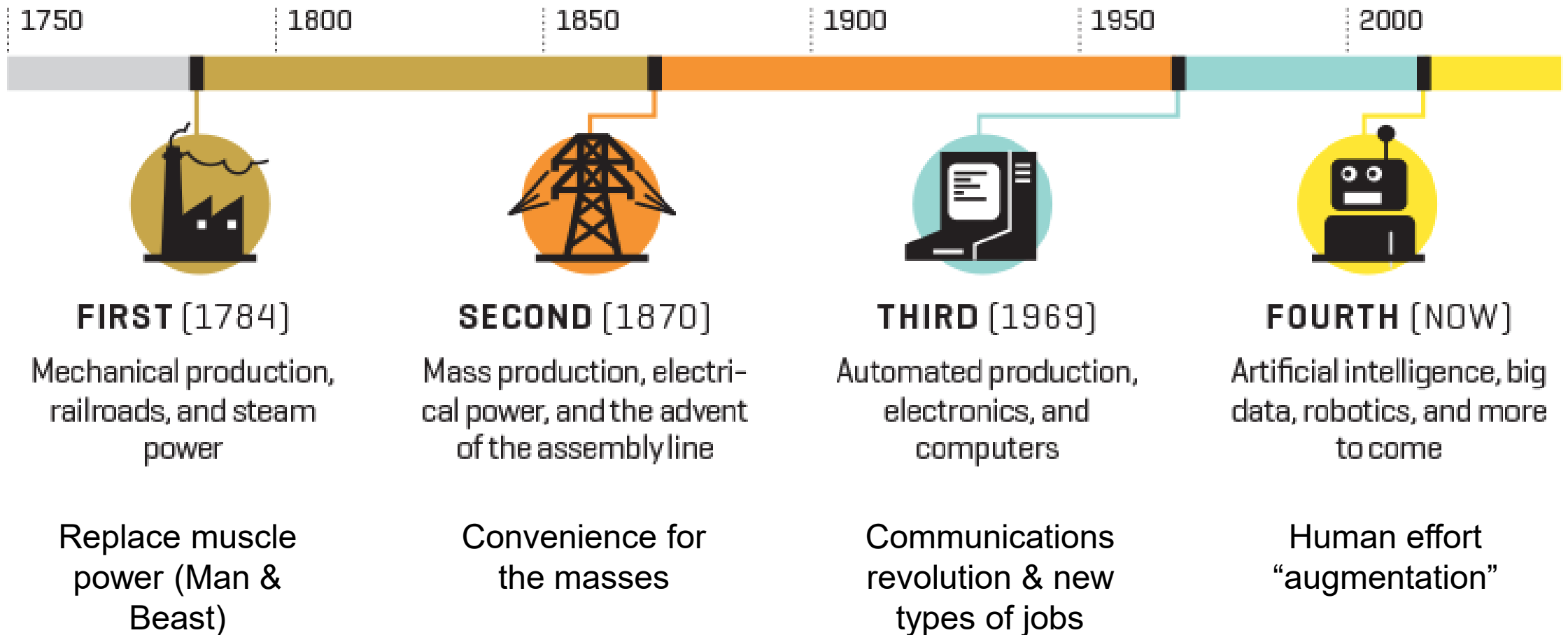
About the Speaker

- Currently Sr. Database/Data Architect @ PayPal
- Has been working with Oracle Databases and UNIX for 3+ decades
- Working on various NoSQL/Big Data technologies for the past 6 years
- Design and Implement High scale systems – Both Oracle and NoSQL
- Author, Technical editor, Oracle ACE *Alumni*, Frequent speaker
- Loves to mentor new speakers and authors!
- <http://www.linkedin.com/in/johnkanagaraj>



Data is the New Oil! (Also “Data is the New Electricity!”)

THE FOUR INDUSTRIAL REVOLUTIONS



Challenge : Explosion

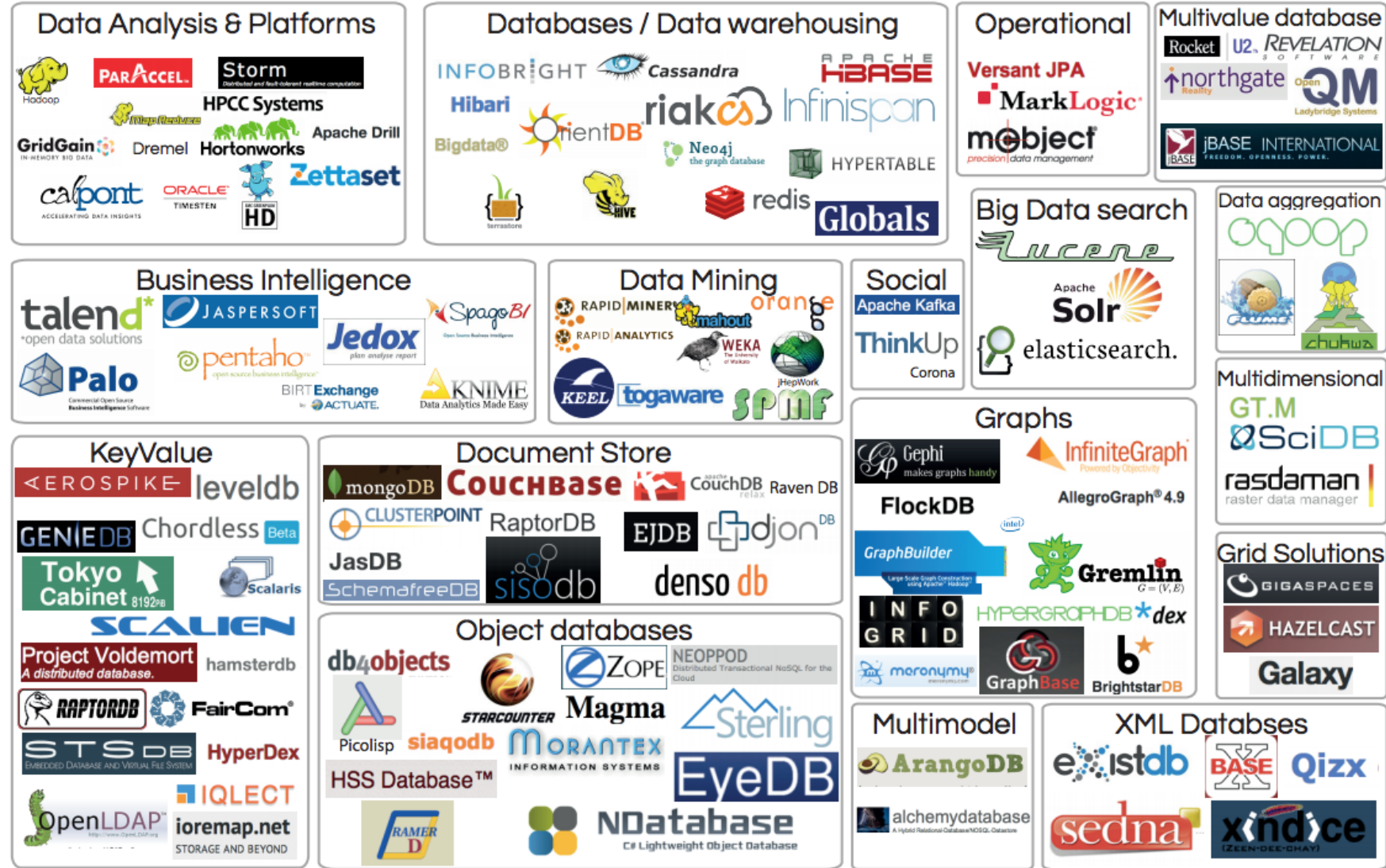
2013

4.4 ZB Digital
Universe
20% Data In Cloud
22% Useful Data

2020

44 ZB Digital
Universe
40% Data In Cloud
37% Useful Data

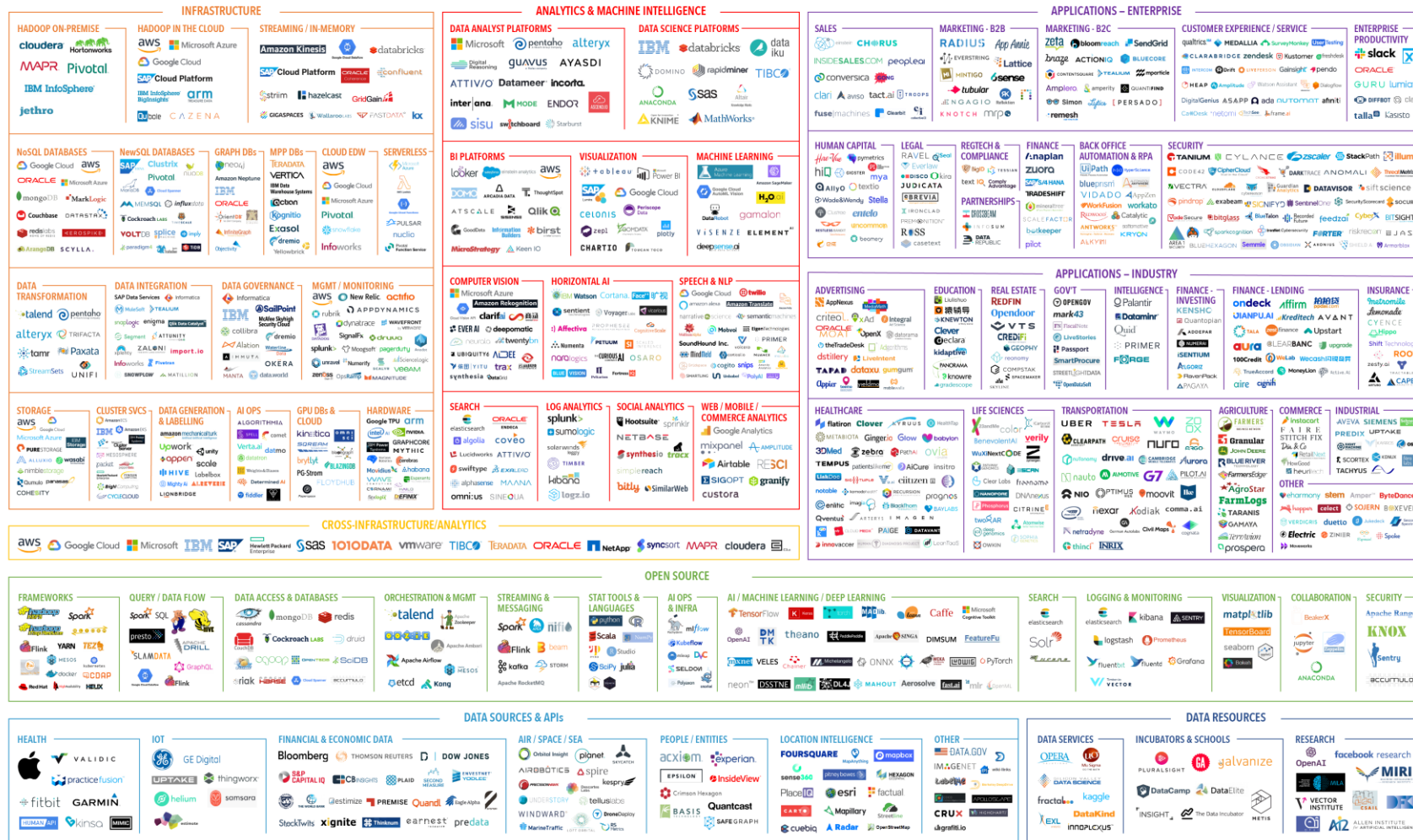
Challenge : Proliferation (2015 Landscape)



Created by: www.bigdata-startups.com

Challenge : Proliferation (2019 Landscape)

DATA & AI LANDSCAPE 2019

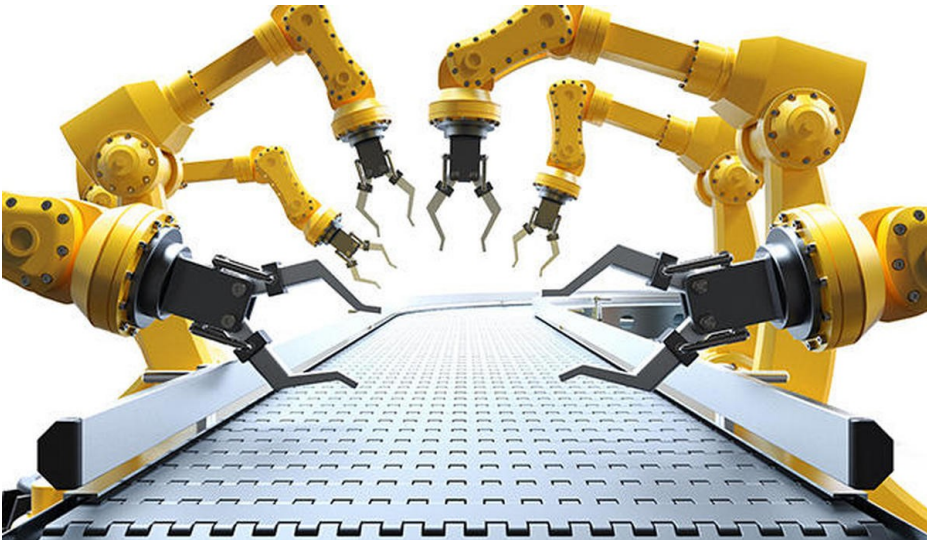
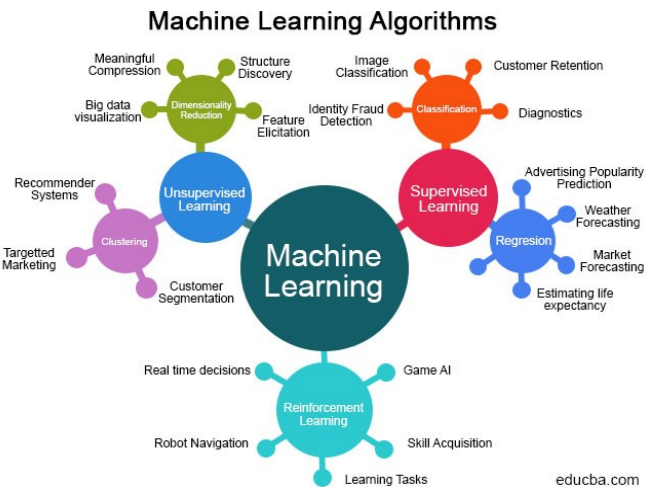
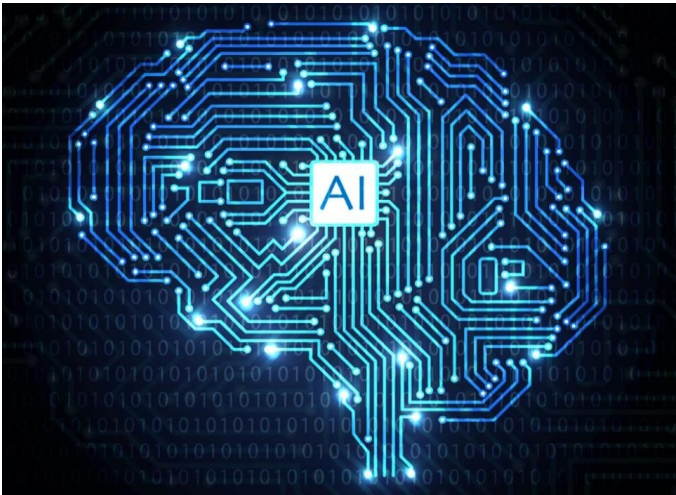


July 16, 2019 - FINAL 2019 VERSION

© Matt Turck (@mattturck), Lisa Xu (@lisaxu92), & FirstMark (@firstmarkcap) mattturck.com/data2019

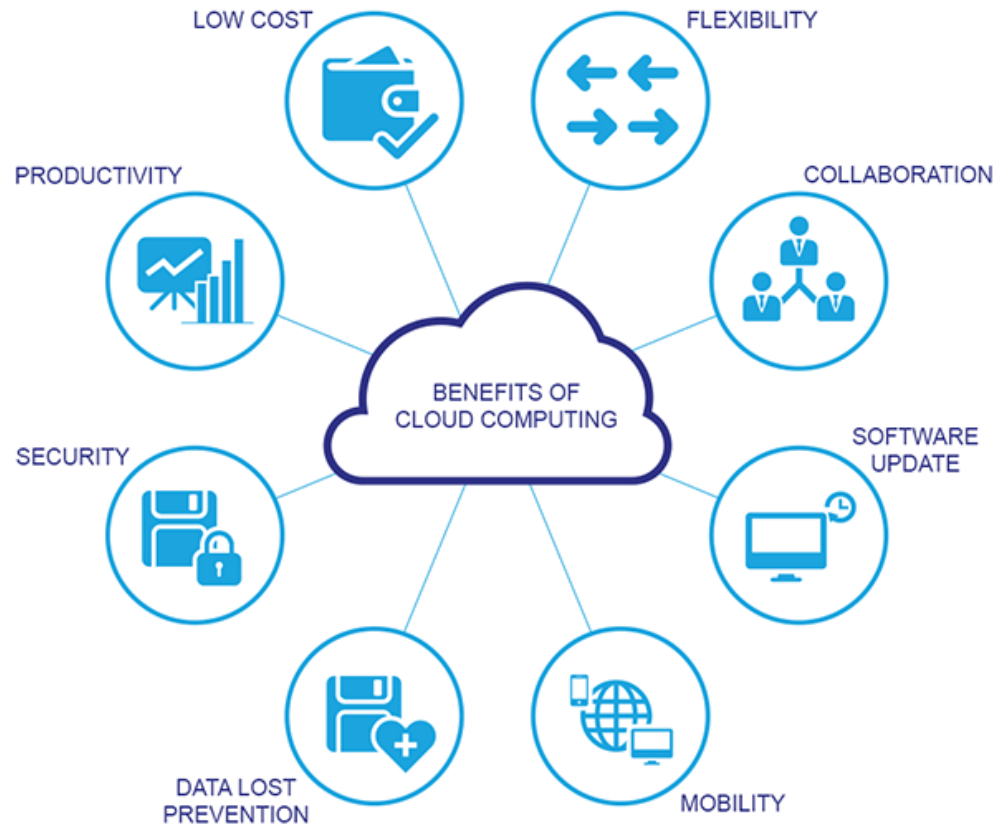
FIRSTMARK 
EARLY STAGE VENTURE CAPITAL

Challenge: Automation

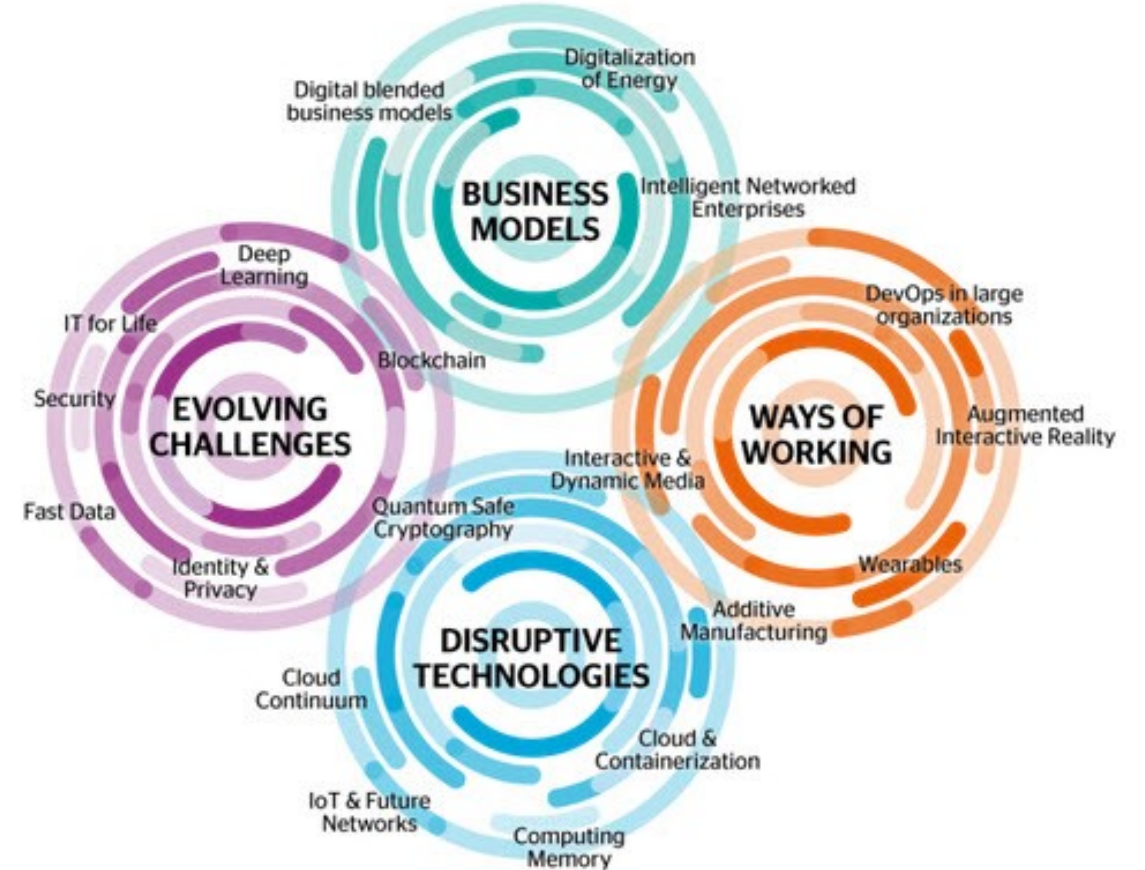


Challenge: Evolution

Cloud computing : CAPEX vs OPEX



Disruptive technologies and new ways of working



Is it different this time around?

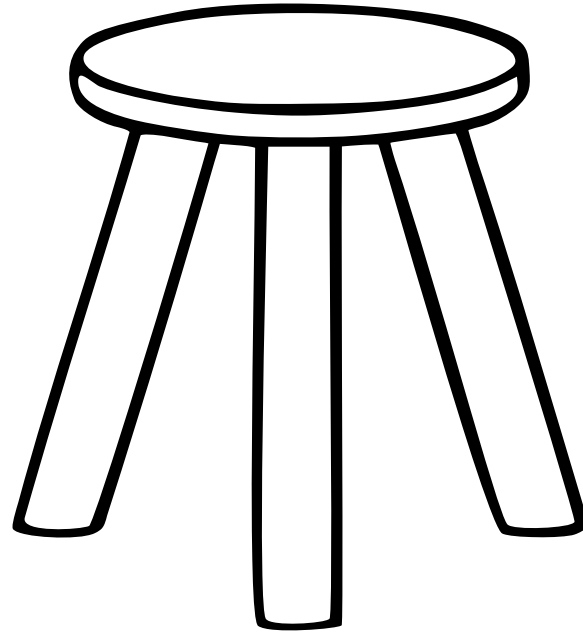
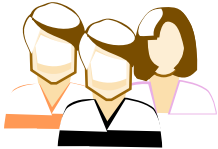
- Artificial *Narrow* Intelligence replacing skilled humans enmasse
 - Self driving Cars and Trucks
 - Internet-of-Things powered by AI based backends
 - Voice powered assistants – Siri, Alexa and Cortana
 - Translation Services – Google Translation
 - Medical Diagnosis and services – Based on IBM Watson
 - “Almost Human” Chat bots augmenting Customer Support
 - RPA – Robotic Process Automation
- Multiple, broad areas undergoing change – Large portions of society affected
- Compressed timelines – Change is too quick to adjust for humans
- Global, Instant access for almost everyone, (mostly w/o national boundaries)
- Learning and advancement not restricted only to formal education

Rising to the Challenge: Tried and tested?!

People, Process, Technology

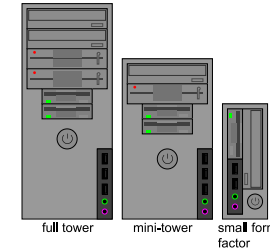
- People

- Recruit or Retrain
- Organize and Lead
- Motivate and Retain
- Manage the culture



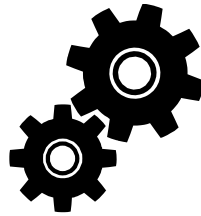
- Technology

- Making the right choice
- Manage the product lifecycle
- Adapt to changes
- Cost management



- Process

- Recognize the need
- Create or Adapt
- Clear Roles and Responsibilities
- Measure, report and adjust



Call to Action

How to get there – For you and for your organization (My view – YMMV!)

You have to change and adapt!

- Know and understand what you are dealing with
- Start your Learning! And continue Learning!
 - Start being curious!
 - **Structured Learning:** LinkedIn Learning, Udemy, Open Courses from Top Universities, Other self-learning paths...
 - **Unstructured Learning:** Blogs, News items, etc.
 - **Compressed Learning:** Boot Camps
 - Apply your learning – Mini projects, Proof of Concepts
 - Operate from your strength! You are already a Data Professional, and probably have a STEM background
 - Persist!
- Join the community
 - Collective, 1-1 accountable learning
 - Contribute and teach someone else!

Your organization needs to change and adapt!

- Define a Business Outcome that AI/ML can enable
 - Problem definition
 - Feasibility study
 - Funding and Support
- Build the Team
 - Internal resources
 - External assistance/consultancy
- Assess the Risk and Reward
 - Risk of AI/ML - governance and bias
 - Measurable gains
- Build the culture
 - The new way of conducting the business – “AI First”
 - Democratized data and AI/ML solutions

