Data & Analytics –
Transforming into a value-driven, convergent space

Anirban Bhattacharya
Practice Head – Data & Analytics – APAC, India (Telecom, Media & Entertainment)
Tech Mahindra

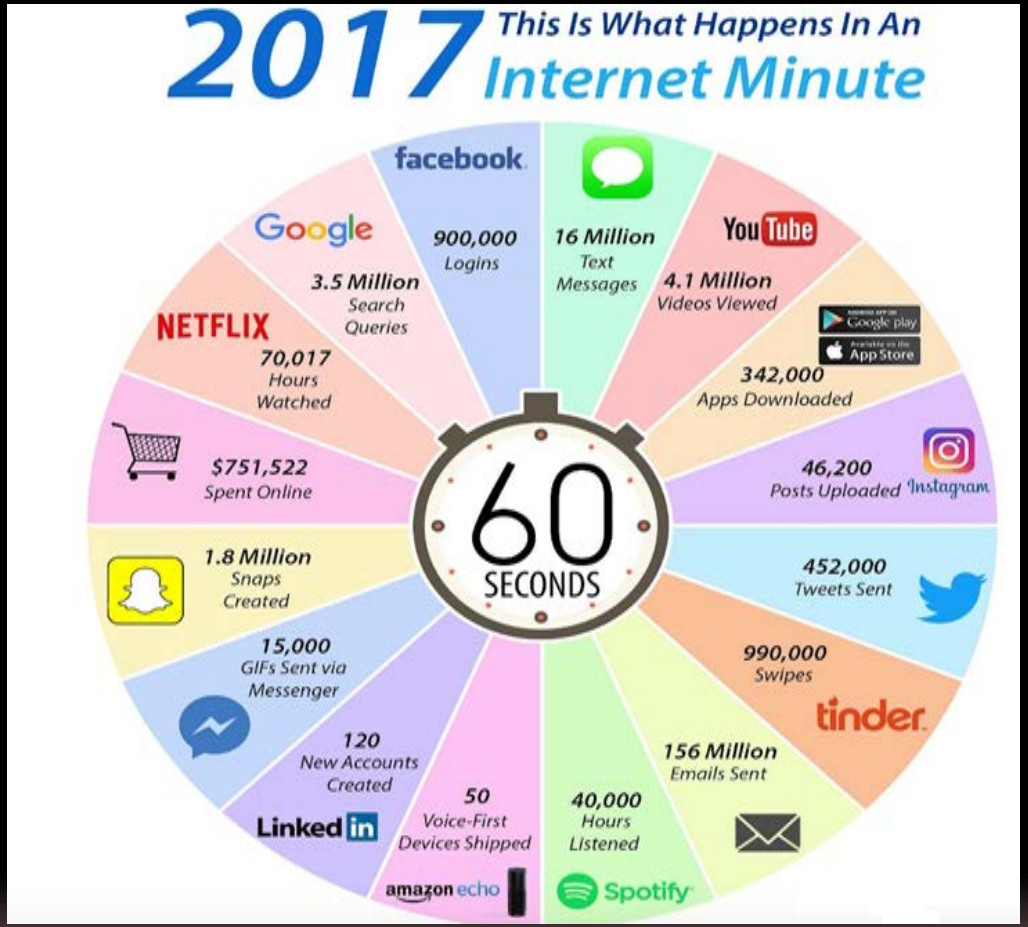
CYBERC SUMMIT 2017
October 2017

I Copyright © 2017 Tech Mahindra. All rights reserved.

Data Explosion – the peril and the opportunity

I Copyright © 2017 Tech Mahindra. All rights reserved.

Just a minute!



That leads us to nothing short of a Data explosion – the reality of today!

Tech Mahindra

43 Trillion
Gigabytes of Data will be created by 2020- 300 times from 2005

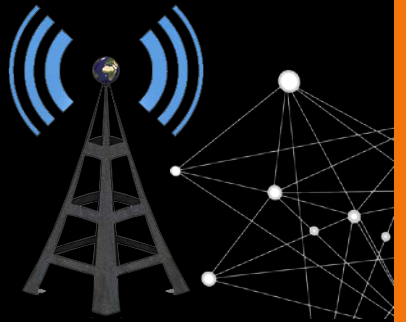
2.3 Trillion
Gigabytes of Data is created everyday

> 6 billion people
have cell phones

Stock exchanges capture
TBs of data
per session

Millions of CDRs, Top-Ups and other events are generated every minute

Almost all modern vehicles have
GPS
enabled



Network connection
per person is going to be >2 for entire world in 2016! Just imaging the population

From Music to medical records, from traffic on roads to social media, our world relies on technology and data- It has to be stored, analyzed and used. These data are huge! And the single term that tries to encompass it is- Big Data.
How and what is it?

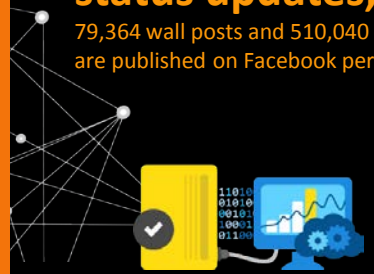
The type of data depends from Organization & industry. For Telecom network providers, we can consider CDRs, Top-Ups, other events, probe data, NoC, site maintenance etc.

Twitter gets
98,000 tweets
per min

600+ videos
are uploaded on YouTube videos per min

695,000 status updates,
79,364 wall posts and 510,040 comments are published on Facebook per min

Google serves more that
94,445 queries
per min



What is the
probability
that the data is > 80% correct?

Poor DQ cost huge loss- US Economy lost
$3.1 Trillion a year

Where has it led us to?

Tech Mahindra

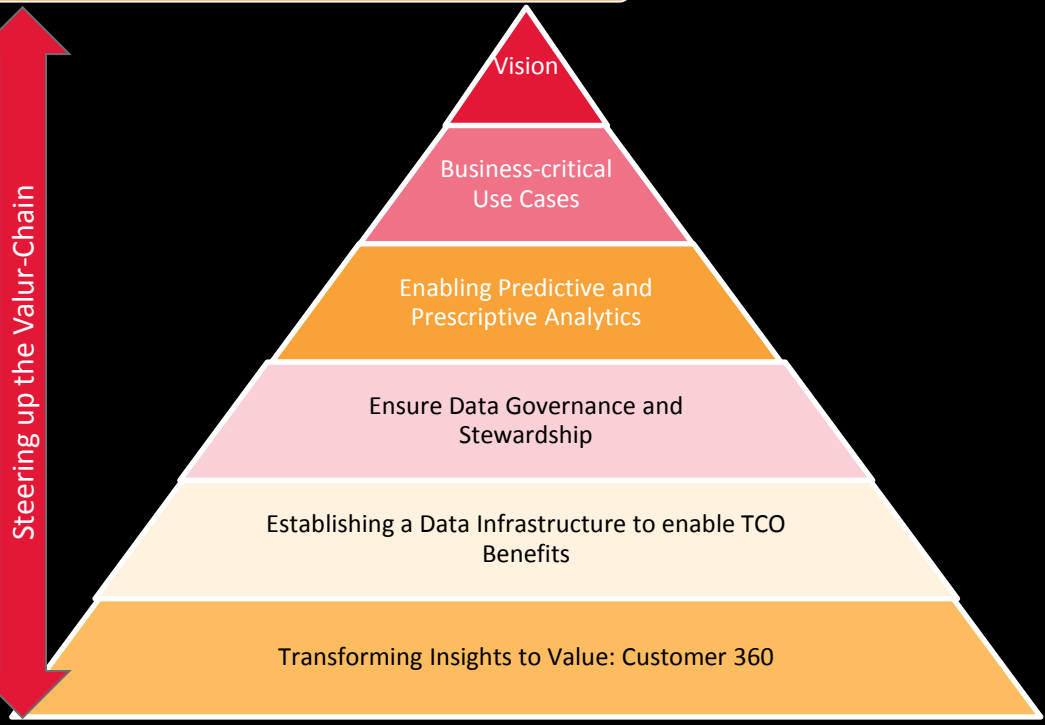
Big Data – a wide canvas

Tech Mahindra

Big Data challenge triggering wide avenues for opportunity and engagement

- Data Driven Transformation through a “Data-First” Culture
- Capability enhancement, enabling Data Driven Decisioning
- Define Platform Arch, establish an Analytics Operating Model
- A Data Governance approach on the unstructured world
- IOT Analytics, Machine Learning
- Enable Digital Transformation leveraging Data & Analytics
- Build Analytics Roadmap for Organization

Steering up the Valur-Chain



The value-chain and a diverse set of technologies

Tech Mahindra

Data Ingestion

informatica
talend
pentaho
TIBCO Jaspersoft
rapidminer
TIBCO StreamBase
hadoop MapReduce
Flume
FTP
kafka

Data Processing & Transformation

Spark
python
splunk>
hadoop MapReduce
Flink
STORM
TRIFACTA
TIBCO StreamBase

Data Analytics

R
WEKA
mahout
Spark
splunk>
sas
SPSS
rapidminer
TIBCO Spotfire
SAP
python

Data Storage

hadoop HDFS
APACHE HBASE
HIVE
Parquet
mongoDB
Cassandra

Data Reporting & Visualization

tableau
TIBCO Spotfire
TIBCO Jaspersoft
MicroStrategy
QlikView
Qlik Sense
Power BI

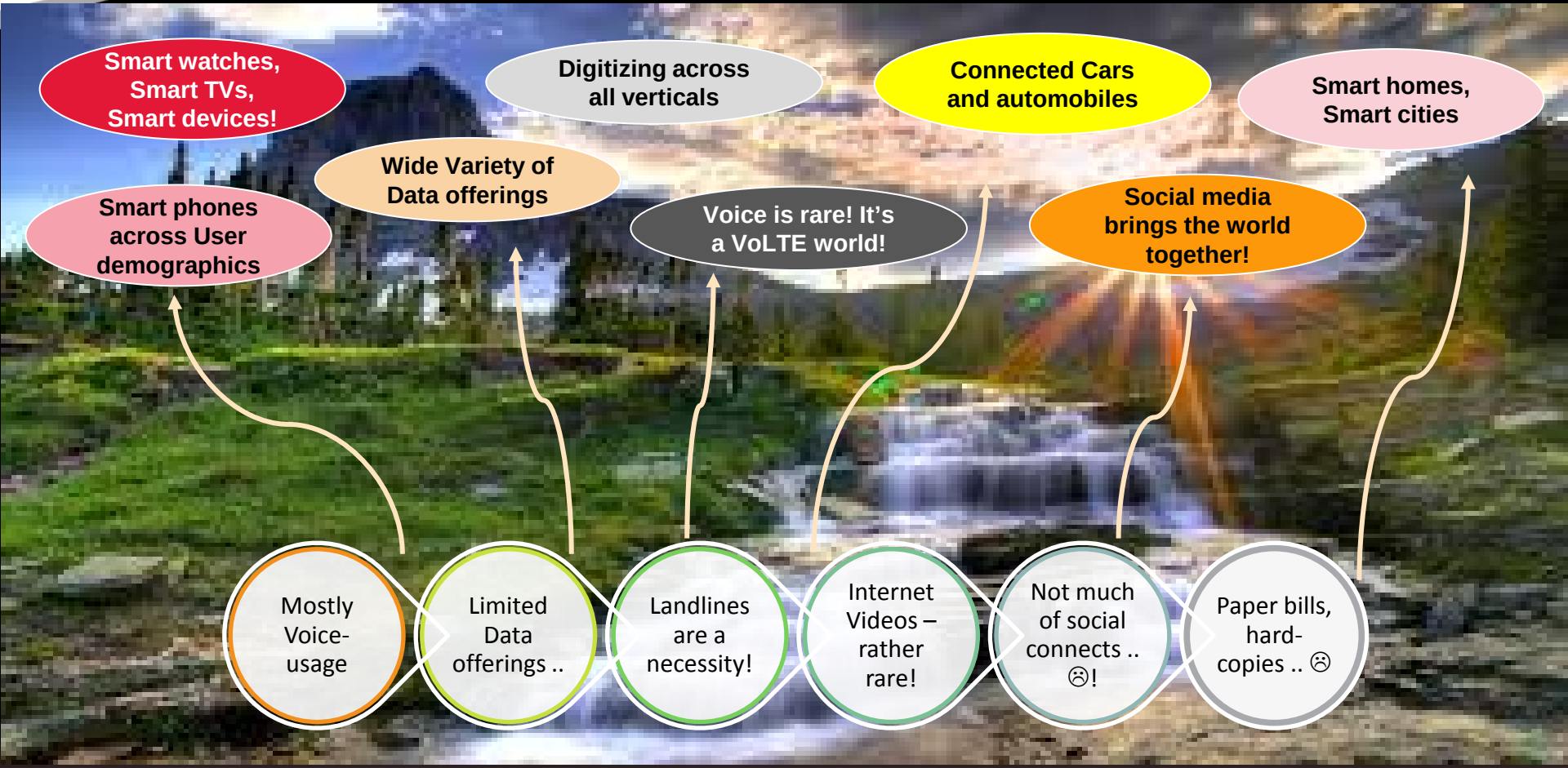
Data Governance & Management

APACHE FALCON

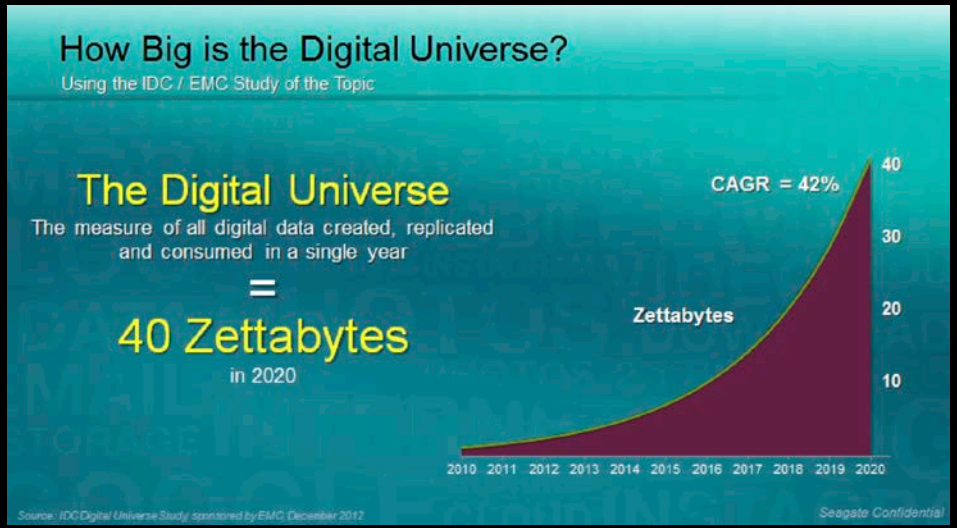
I Copyright © 2017 Tech Mahindra. All rights reserved.

The shift is significant!

Tech Mahindra



Birth of the digital paradigm



New trends in the Telecom space

Tech Mahindra

While they are still relatively niche products, wearables such as smart watches and fitness bands have seen tremendous percentage growth.

The consumer-oriented "things" that comprise the IoT—including wearables, connected cars, smart homes (e.g., lighting, security, entertainment), and the government and enterprise-connected "things" such as smart businesses (e.g., fleet management), and smart cities (e.g., parking, city lighting, asset monitoring and tracking, and video security)

Amazon style platform to support an a la carte approach to selling video programming, potentially from a wide range of sources delivering content to any screen is finally becoming a reality, enabled by advances in network technology and higher speeds, as well as enhanced content at the carrier level, whether owned or resold

Smartphone sales are still strong, the highest growth percentages coming in the 45-54 and 55+ age demographics–groups that have previously lagged behind younger consumers due to Mobile video calling and other multimedia offers

Modernizing Telco Operations with help of Digital Transformation having a huge scope for Automation Cater to incremental demand for Value-added Infrastructure Services

the autonomous vehicle has quickly emerged as a viable and highly desired product for consumers

Telcos enabling Digital transformation for other verticals

Tech Mahindra

Healthcare

Digital Medicine

Digital Genomics

AI Physicians

Mobile Care

Entertainment

Movie Tickets

Concert/Shows

Sports

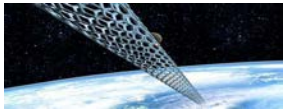
Content

Transportation

Self Driving Vehicles

Robotic Drones

Space Elevator



Manufacturing

3D Printing

Robotics

Nano Manufacturing

Banking

Crypto Currencies

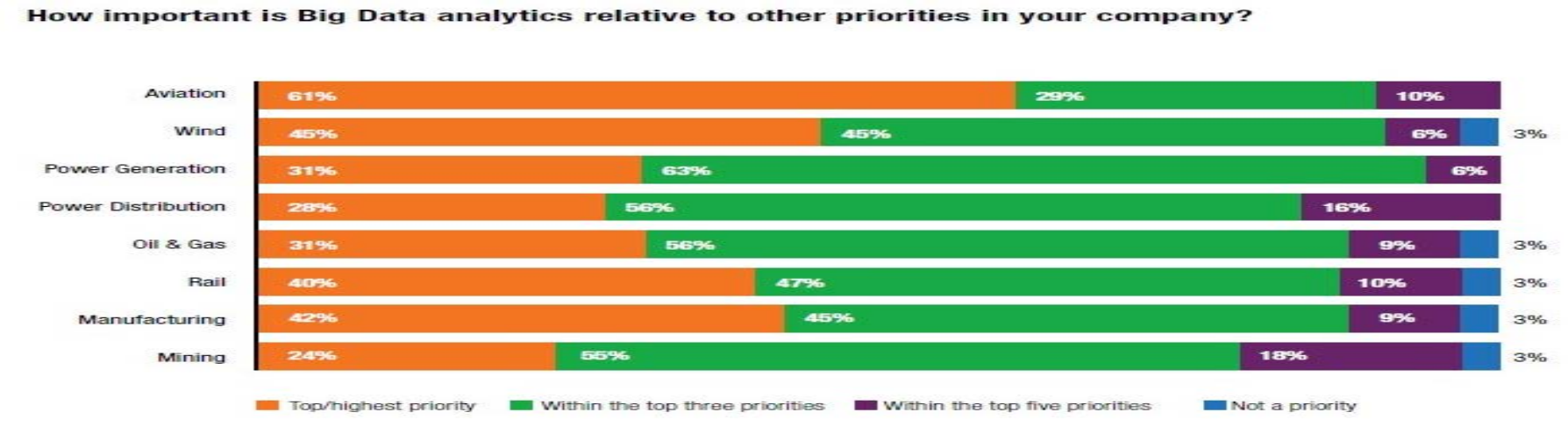
Biometric Authentication

Crowd Funding

Virtual Bank

Telco is a key enabler to other industries for their Digital Transformation through its Digital products & service offerings and differentiated Business Models based on strategic partnerships

A few expected fallouts



Big Data market is expected to grow from USD 28.65 Billion in 2016 to USD 66.79 Billion by 2021, at a high Compound Annual Growth Rate (CAGR) of 18.45%

Spending on Self-Service Visual Discovery and Data Preparation Market Will Grow 2.5x Faster Than Traditional IT-Controlled Tools

Spending on Cloud-Based BDA Technology Will Grow 4.5x Faster Than Spending for On-Premises Solutions

Big Data, Data Science & Analytics Professional Services will Have a CAGR of 23%

I Copyright © 2017 Tech Mahindra. All rights reserved.

What lies beneath and what is sought for – the Big Data asks of today!

What is expected!

A platform integrating internal and external business information

Powered by Self-service & ad-hoc analysis

Predictive Analytics enabling strategic decisions

Enabling effective Campaign Management

Scalable, extensible, modular and agile

Offering a single-version-of-truth of the Customer

Enabling Digital Analytics for the future

Today's business realities and challenges

Contextual Marketing

Customer Experience

IOT, M2M

Deriving and monetizing value from data

Social Grouping & Networking

The TCO challenge with the current data explosion

Data and Video – Voice is "legacy"

Social Media & Sentiment Analytics

Diverse market and consumer behavioral trends

Data Privacy – a growing urgency

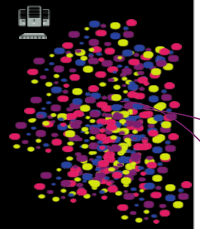
Big Data and the 4 Vs

Tech Mahindra

Big data in reality

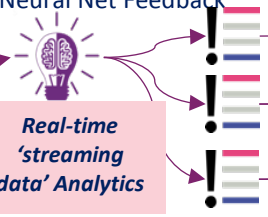
Volume + Variety + Velocity + Value

Filter	Detect	Act	Diagnose	Restore
Technique automatically aggregates, correlates and then ranks the "Significance" of each Event Type	Anomaly Detection Techniques • Textual categorization/Time • Textual similarity and Topology, Knowledge and Recipe • Neural Net Feedback	Real time notification by automatic incident creation by integrating with CRM and incident management.	Using past records, history, create diagnosis for every incident	Real-Time action triggering



Events

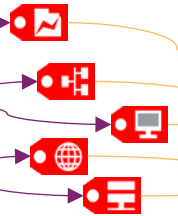
Events: Millions of Events, CDRs Get Triggered from OTT Apps, Services & Networks



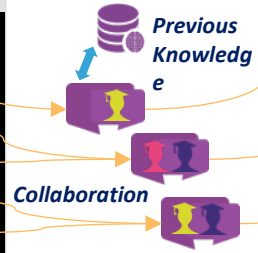
Detect Anomalies

Benefits:

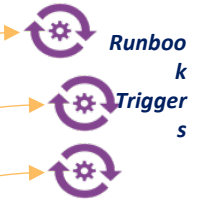
- *>80 % reduction of actionable items*
- *Diagnostic Alert narrative*



Actionable Intelligence



Prescriptive Analytics



Automation

Combating the 'V' challenges

Tech Mahindra

Volume
- Network 4g Data storage
- Mobile Data Explosion
- Data increase from xTB to 100xTB
- N/W Alarms and data growing exponentially

+

Velocity
- Real time alerts
- Real time analytics
- Zero latency reporting

+

Variety
- Alerts and Logs
- Sensors and Devices
- Machine-generated communications and transactional activity

=

Challenge
- Increased cost of Operation
- Requirement of diverse Technology array
- Ecosystem maturity
- Demand for new age requirements/Use cases

Data Sources

Multi-machine Commodity Hardware

Big Data platform powered by hadoop

Data Landing

Data Lake

Data Virtualization

Semantic Layer

- Flexible and resilient
- Immensely Scalable
- Faster processing and movement
- Huge TCO Improvement
- Improved Decision-enablement
- Significantly Reduced Time-to-Insight

And delivering business value across the chain

Tech Mahindra

- Portfolio & Contextual Campaigning
- Dynamic Customer Segmentation
- Customer Experience improvement
- Omni-channel experience Analytics
- Data Anonymization
- Churn Propensity Analytics
- Enabling an effective Product strategy
- Network Congestion Control
- Call Drop Analytics
- Enabling Revenue Assurance & Fraud Prevention
- Voice of Customer 360°
- Social Media Analytics
- Network Experience Improvement
- Next-Best Action / Next-Best Offer

I Copyright © 2017 Tech Mahindra. All rights reserved.

Big Data – the Data-value realization challenge

Tech Mahindra

Read & visualize data from Data Lake

Join & transform data from multiple sources

Generate reports from data on HDFS, Hive, RDBMS

Reduce Efforts for Techie, IT & Data Scientist for data insights?

Volume ✓

Velocity ✓

Variety ✓

Value

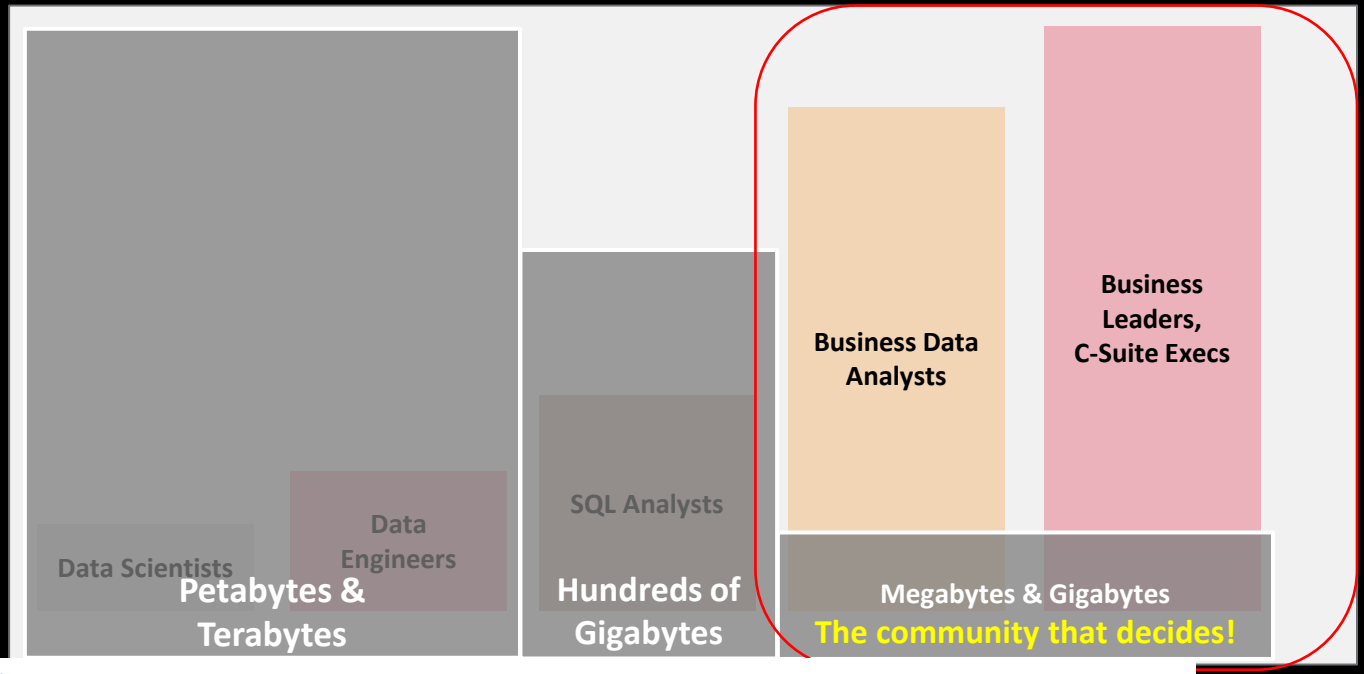
Data is available, BUT….

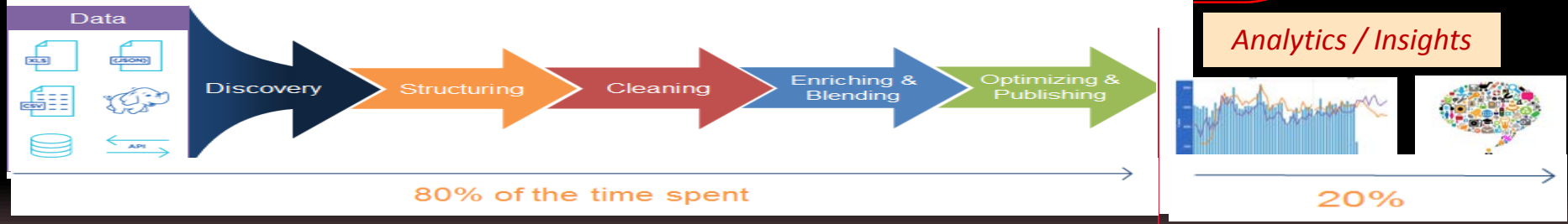
Can I access and decipher it? **AND**

Can I derive value from it?

Data access trends across user-communities – the 80:20 paradox

Tech Mahindra





Addressing the pain
.. accelerating the realization of Data-value

Tech Mahindra

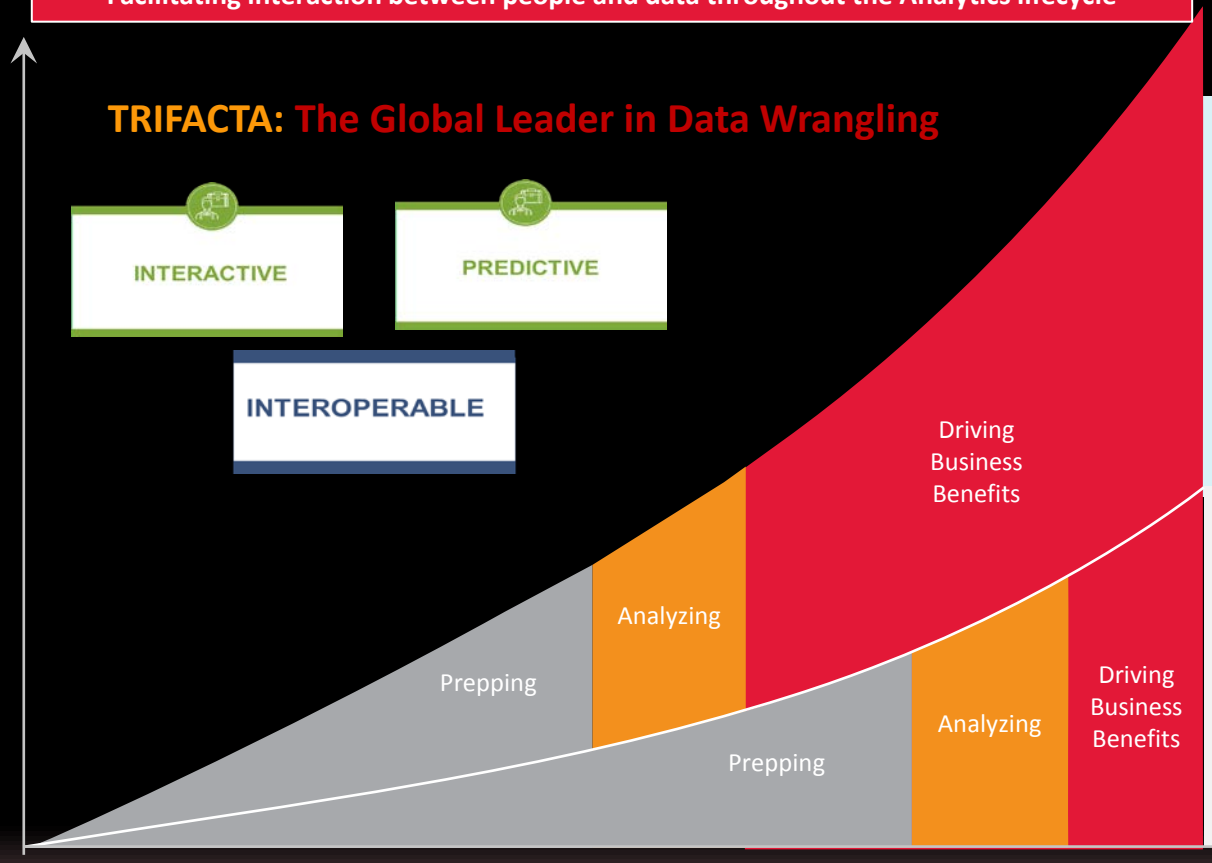
Facilitating interaction between people and data throughout the Analytics lifecycle

TRIFACTA: The Global Leader in Data Wrangling

INTERACTIVE

PREDICTIVE

INTEROPERABLE



TRIFACTA

Excel

IT

DIY (Scripting)

Reports / BI

Analytics / Insights

BTL Campaigns

Acknowledged as the leader across the board

Tech Mahindra

No. 1 by Customers

PEPSICO, MarketShare, GoPro, NASA, ebay, Lockheed Martin, Deloitte., CITCO, dish, orange, UnitedHealth Group, Bell, Moody's, BCG, CDC, SANOFI, Kaiser Permanente, ENSTAR, Munich RE, Juniper, RBS, gsk

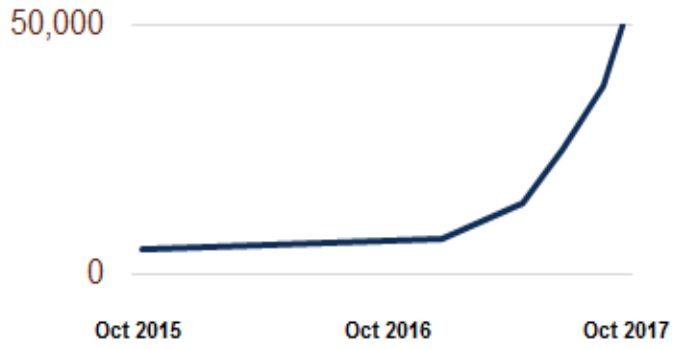
No. 1 by Analysts

Dresner Advisory Services	Dresner Advisory Services	Bloor	F
2015	2016	2016	2017
#1 End User Data Preparation Vendor		#1 Self-Service Data Prep & Cataloguing	#1 Data Preparation Tool

No. 1 by Partners

Google Cloud Platform, cloudera, tableau, Qlik, amazon web services, slalom, tamr, accenture, HORTONWORKS, ARCADIA DATA, BARDESS, ZOOMDATA, Microsoft Azure, WATERLINE DATA, atscale, MAPR

No. 1 by Users



Moving towards a ‘converged’ BI landscape

I Copyright © 2017 Tech Mahindra. All rights reserved.

Huge storage and rapid Processing – how have we done?

Tech Mahindra

Storage capacities of hard drives have increased massively over the years (Moore's law), but access speeds—the rate at which data can be read from drives (or written to drives) — have not kept up *in comparison to Storage Capacity.*

- For Example:
  - One typical drive from 1990 could store 1,370 MB of data and had a typical transfer speed of 4.4 MB/s, so you could read all the data from a full drive in around five minutes.
  - 20 years later, one terabyte drives are the norm, but the transfer speed is around 100 MB/s, so it takes more than two and a half hours to read all the data off the disk.

Unfortunately, a rather long time to READ all data on a single drive—and writing is even slower!!

Solution? Obviously Parallel Processing!!

- Most of these models use commodity hardware and hence were developed to address the issues of hardware failure and combining input / output from multiple discs. The most popular one is Hadoop as on date.
  - **The distributed storage is provided by HDFS**
  - **The analysis is provided by Map-Reduce.**
  - **It has a Distributed Data Storage and Distributed Data Processing Framework that handles petabytes of Data in limited time**

In April 2008, Hadoop broke a world record to become the fastest system to sort a terabyte of data - Running on a 910 node cluster, Hadoop sorted one terabyte in 209 seconds!!

Increasing significance of 'analysis' across the business-chain

Tech Mahindra

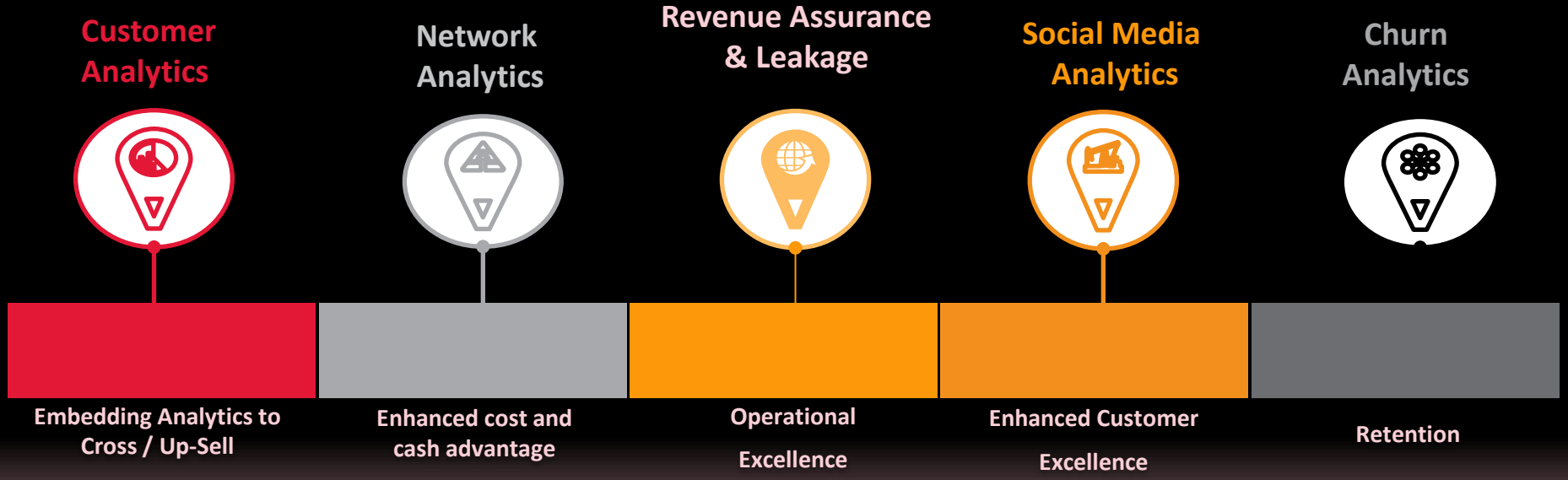
Technology focused offerings with focus on enabling our clients to become *Insights-Powered, Data - Driven Enterprises*

Using some of the latest technologies & Platforms With an intention to drive value with *'Cultivating advanced Analytics'* theme

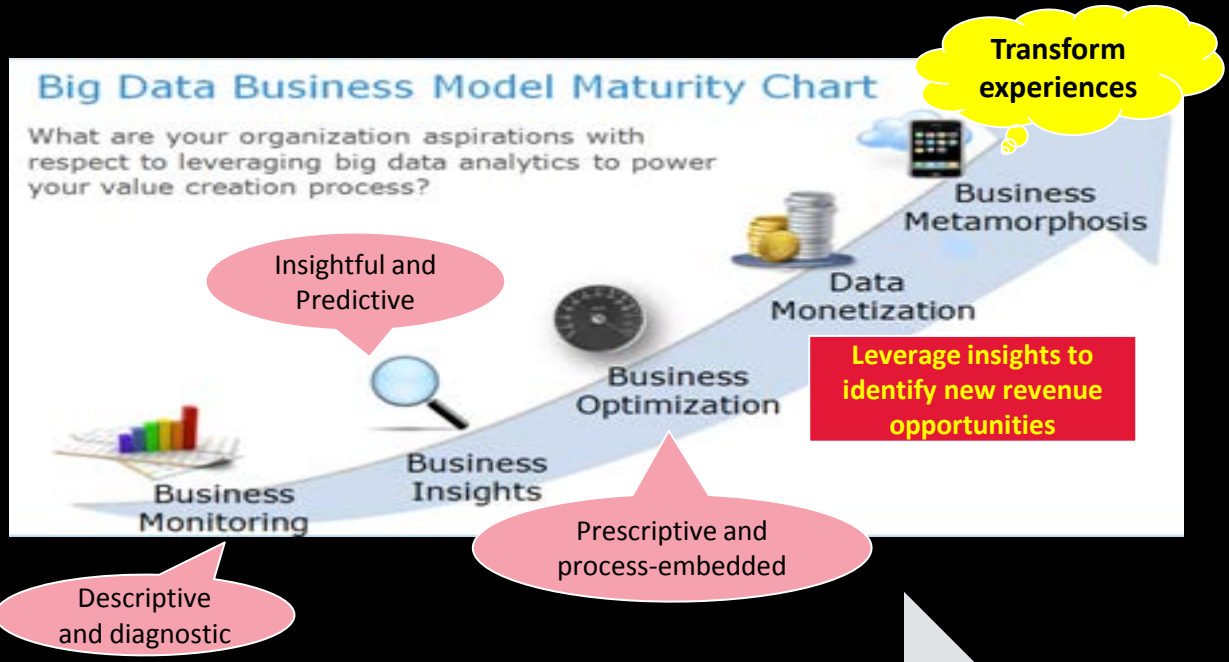
Not just Technology Solutions but .. .. comprehensive Business Solutions

Business focused offerings with specialized and pointed use-cases for customers across various process streams

Focus is to deliver business value harvesting our *experiences, domain knowledge and Platforms*



Lots of data! .. but what is it's value?



Assessment and availability

Storage and segregation

Insights and intelligence

Segment and sell

How could the data be 'monetized'?

Tech Mahindra

The *mantra* – Keep Your Money and **just give me your DATA!**

Internal Monetization

Companies using customer's data insight for cross/up sales of their product & services to their customers

External Monetization

Organizations can monetize customer's data by collaborating with a range of customer facing industries like Advertising, Marketing, Financial Services, Retail, Ecommerce, etc.

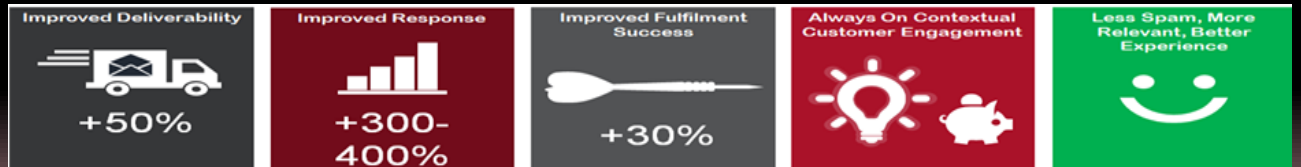
An effective Data Monetization solution – 2 imperatives…

Capture diverse info, structured and unstructured alike

Solution to capture and process all forms of data featuring across various sources

Prevent Security breach of any Customer PII

A robust and automated Data-masking and anonymization solution



No more merely the eyes and ears of the enterprise but beyond … possibly its brains and hearts?

Tech Mahindra

Highly Customer-Centric

Advanced Analytics

M2M & IOT

Technology Transformation

Cognitive and AI

Hybrid BI platform

Business-Model re-engineering

Gain Share Business Model

Data Management Services

- Packaged BI solutions
- BI Modernization
- Analytics as a Service

Crowd-sourced from community

Engage third-party/ niche players

Engage strategic partners

Partner with Academia

Customer Co-innovation

Pro-actively assess and address Customer Needs

The world has gone F A R ahead!

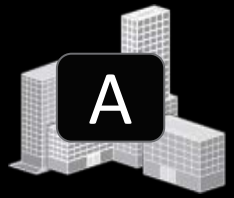
Tech Mahindra

Lets **'Dimensionalize'**, lets build a Star-schema

Can we flatten the Star further?

Columnar storage – they are good!

Could we please think Schema-less!!



- Lets refresh once a day – we should be good!

The "Daily batch" era

Multi-batch per day norm

- Could we get a feed every 8 or 10 hours? 24 hours is long!

- 8 hours .. Oops! Need it hourly!!

Latency steeps up!

Near real-time days

- Oh no!!! Would need a data stream – delays proving costly

- Can I have analytics embedded in my process!!

Intelligence into the transactions

E-Mail Conversation between Tom and Julie

Instant chats as if Jill and Bob were speaking!

A group discussion – all views invited!

Not restricted to humans only!

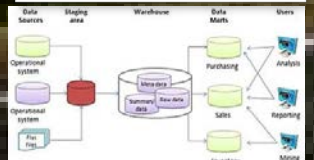
Machines are also gaining intelligence!!

A long-journey into today's truly 'Converged' world

Tech Mahindra

Consistent!

EDW, single-version-of-truth, consistent, conformed



Conjoint ..☺

Advent of the Appliance – a con-joint approach of putting the S/W and H/W in step

Cloudy ..!

Focus on the Cloud – Public, Private, Hybrid, Community

Conglomerated

Genesis of Hybrid BI – Classical RDBMS complemented with a Hadoop platform

Converged!!

Bringing together the laptops and mobiles; interchangeability between products and services; thinning down the lines between the 'Analytical' and 'Transactional' worlds through 'embedded' Analytics into transactional Business processes

Thank you

For more information, please visit www.techmahindra.com

Disclaimer

Tech Mahindra Limited, herein referred to as TechM provide a wide array of presentations and reports, with the contributions of various professionals. These presentations and reports are for information purposes and private circulation only and do not constitute an offer to buy or sell any services mentioned therein. They do not purport to be a complete description of the market conditions or developments referred to in the material. While utmost care has been taken in preparing the above, we claim no responsibility for their accuracy. We shall not be liable for any direct or indirect losses arising from the use thereof and the viewers are requested to use the information contained herein at their own risk. These presentations and reports should not be reproduced, re-circulated, published in any media, website or otherwise, in any form or manner, in part or as a whole, without the express consent in writing of TechM or its subsidiaries. Any unauthorized use, disclosure or public dissemination of information contained herein is prohibited. Individual situations and local practices and standards may vary, so viewers and others utilizing information contained within a presentation are free to adopt differing standards and approaches as they see fit. You may not repackage or sell the presentation. Products and names mentioned in materials or presentations are the property of their respective owners and the mention of them does not constitute an endorsement by TechM. Information contained in a presentation hosted or promoted by TechM is provided “as is” without warranty of any kind, either expressed or implied, including any warranty of merchantability or fitness for a particular purpose. TechM assumes no liability or responsibility for the contents of a presentation or the opinions expressed by the presenters. All expressions of opinion are subject to change without notice.

 Copyright © 2017 Tech Mahindra. All rights reserved.