

# Digital Transformation

## Class # 3: Technology Trend # 2 – Big Data & Artificial Intelligence (AI)

EPITA | Fall 2021

---

Valeriu Petrulian

# Digital Transformation Class 3



- Admin
- Recap
- Thoughts for the day

# Course Breakdown

## EPITA Fall 2021

Class	Date & Time   Topics
Class # 1	Welcome to the Digital Economy!
Class # 2	Technology Trend # 1: Internet of Things (IoT), Blockchain
Class # 3	Technology Trend # 2: Big Data & Artificial Intelligence (AI)
Class # 4	Digital Platforms
Class # 5	Final Presentations and Course Wrap-Up

### Today's Readings:

- **Eric Colson.** *What AI-Driven Decision Making Looks Like*. Harvard Business Review, July 08, 2019 issue
- **Andrew McAfee, Erik Brynjolfsson.** *Big Data: The Management Revolution*. Harvard Business Review, October 2012



# Big Data & Artificial Intelligence (AI)



- Data is everywhere, in large volumes (hence, Big), but where does it come from? And why has it become so important?
- Artificial Intelligence (AI) is a predominant digital technology
- Together, as fundamental Digital Transformation trends, they have the potential to transform radically the way we interact, work and, possibly, think



# Big Data & AI | Introductory considerations



IDC Digital Universe 2014 Report.  
<http://www.emc.com/collateral/analyst-reports/idc-digital-universe-2014.pdf>

# Artificial Intelligence (AI) | Fundamentals

## ARTIFICIAL

- “Made or produced by human beings rather than occurring naturally, especially as a copy of something natural. Not natural.”
- Unreal
- Factitious
- Fabricated

## INTELLIGENCE

- « The ability to acquire and apply knowledge and skills.”
- Brilliance
- Intellect
- Judgment
- Perception

# Artificial Intelligence (AI) | Fundamentals

## Milestones

- 1950 – Turing Test
- 1956 – Dartmouth College conference, term is coined
- 1965 – ELIZA natural language software
- 1997 – Deep Blue defeats chess world champion Gary Kasparov
- 2009 – Google self-driving car prototype
- 2011 – IBM Watson wins at Jeopardy!
- 2014 – Personal assistants
- 2016 – AlphaGo victory
- 2018 – Turing Prize awarded to Deep Learning pioneers Y Bengio, G Hinton, Y Le Cun



# AI

## A First Classification

- « **Artificial Intelligence** » = a set of several techniques, each one of which has more precise objectives than just « intelligent reasoning » :
- (1) Systems that think like humans (e.g., cognitive architectures and neural networks);
- (2) Systems that act like humans (e.g., pass the Turing test via natural language processing; knowledge representation, automated reasoning, and learning),
- (3) Systems that think rationally (e.g., logic solvers, inference, and optimization);
- (4) Systems that act rationally (e.g., intelligent software agents and embodied robots that achieve goals via perception, planning, reasoning, learning, communicating, decision-making, and acting).

**Source:** White House Report. *Preparing for the future of artificial intelligence*. October 2016

# AI

## A Second Classification

- **Perception** (images, sounds, forms), ex: machine vision
- **Natural Language Processing**
- **Planning and Navigation** (problem solving for operational research)
- **Knowledge representation**
- **Logical reasoning** (symbolic processing, expert systems)

**Source:** White House Report. *Preparing for the future of artificial intelligence*.  
October 2016

# AI

## A Third Classification

- (1) “Symbolists” use logical reasoning based on abstract symbols,
- (2) “Connectionists” build structures inspired by the human brain;
- (3) “Evolutionaries” use methods inspired by Darwinian evolution;
- (4) “Bayesians” use probabilistic inference; and
- (5) “Analogizers” extrapolate from similar cases seen previously

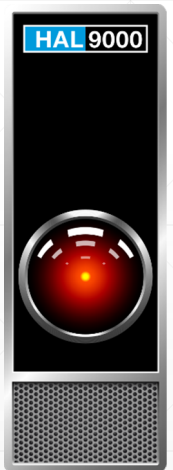
**Source:** White House Report. *Preparing for the future of artificial intelligence*. October 2016

# AI Fundamentals

## Combining several AI techniques



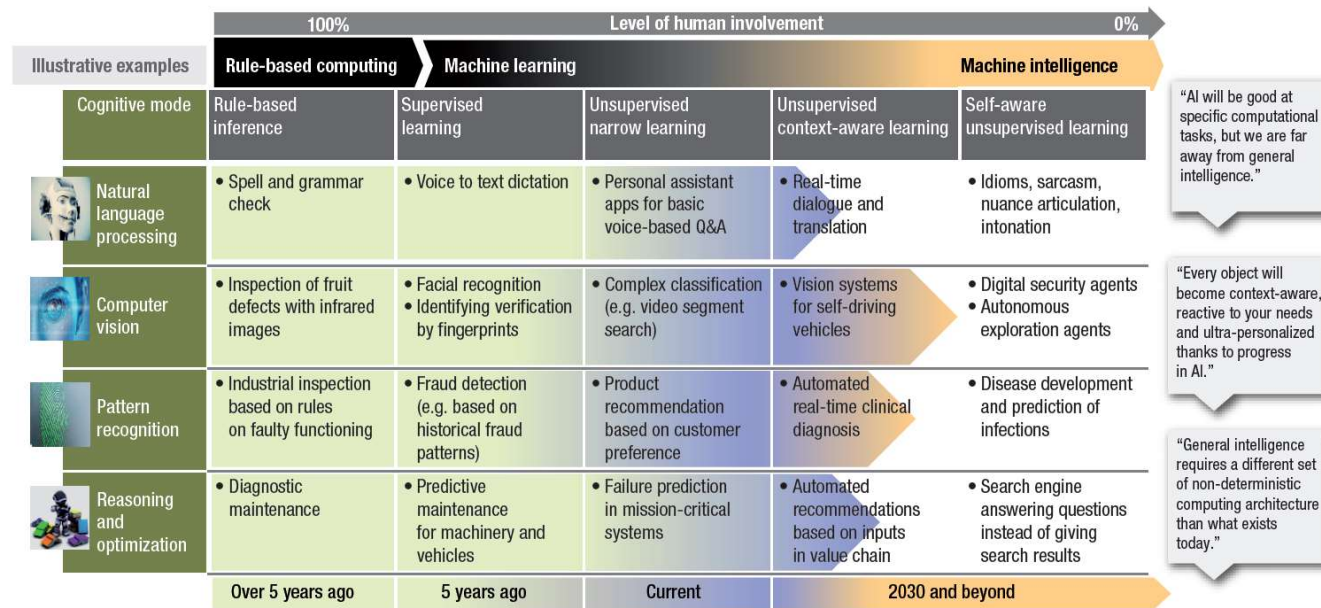
- « General Problem Solver » is a program created in 1959 by Herbert Simon, Alan Newell and Cliff Shaw that implements the idea of complex problem resolution
  - It has generalized the principle that, as soon as there is a need to solve problems beyond the limited scope of a given AI technique, then several AI techniques need to be combined
  - Current examples include IBM's Watson platform, as well as similar platforms from other technology companies (Einstein at salesforce, Leonardo at SAP, etc...)
- More generally, it allows to create a space between **Artificial Narrow Intelligence (ANI)** and **Artificial General Intelligence (AGI)**



All images source: Wikipedia

# Artificial Intelligence State of the Art

Figure 7: Development of AI and its future state



Sources: Company websites; A.T. Kearney; A.T. Kearney/World Economic Forum workshop, November 2016; expert interviews

World Economic Forum, in collaboration with AT Kearney. *Technology and Innovation for the Future of Production: Accelerating Value Creation*. WEF White Paper, Geneva, March 2017.



# Class Discussion

## GOFAI

- Previous form of « intelligent systems »
- « Expert systems », rules-based
- They have been around for some decades
- Widespread adoption by business organizations

## New AI FORMS (ML, DL, ...)

- Recent forms of AI
- Adoption is influenced both by their novelty and by pre-existing « intelligent systems »
- History matters?
- Paradigm shift?

**IS THEIR COEXISTENCE AND/OR COMBINATION POSSIBLE?**



# Artificial Intelligence

## Machine Learning

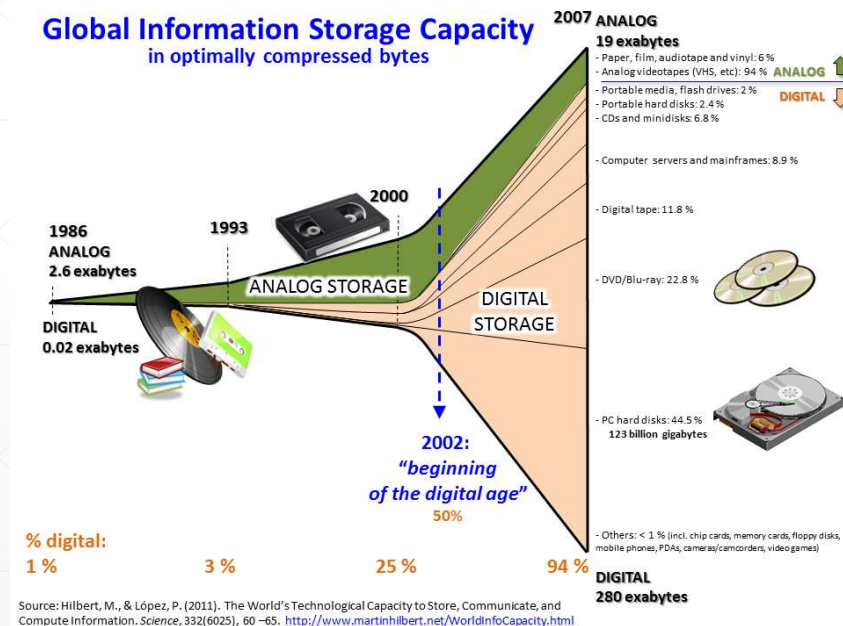
- Wikipedia – « **Machine learning (ML)** is the scientific study of algorithms and statistical models that computer systems use to perform a specific task without using explicit instructions, relying on patterns and inference instead.”
- Learning methods:
  - Supervised learning, Unsupervised learning, Reinforcement learning, Self learning, Feature learning, Sparse dictionary learning, Anomaly detection, Association rules
- Training DATA, Test DATA

# Artificial Intelligence

## Deep Learning

- Wikipedia – « **Deep learning** (also known as **deep structured learning** or **differential programming**) is part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.”
- DL relies on Artificial Neural Networks (ANNs) and Deep Neural Networks (DNNs)
- DL and ML have in common the use of LARGE DATA SETS

# Big Data Fundamentals



**Big data** is a field that treats ways to analyze, systematically extract information from, or otherwise deal with data sets that are too large or complex to be dealt with by traditional data-processing application software. Big data was originally associated with three key concepts: *volume*, *variety*, and *velocity*.

[Source: [https://en.wikipedia.org/wiki/Big\\_data](https://en.wikipedia.org/wiki/Big_data) [Image File:Hilbert InfoGrowth.png|thumb|Hilbert InfoGrowth]]

# Big Data | Fundamentals

## Big or Smart?

- Definitions

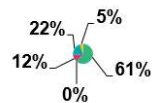
- Oxford English Dictionary: « **“big data** *n.* Computing (also with capital initials) data of a very large size, typically to the extent that its manipulation and management present significant logistical challenges; (also) the branch of computing involving such data.”
- <https://whatis.techtarget.com/definition/smart-data>: **“Smart data** is digital information that is formatted so it can be acted upon at the collection point before being sent to a downstream analytics platform for further **data** consolidation and analytics.»

# Today's Digital Economy = Data Economy

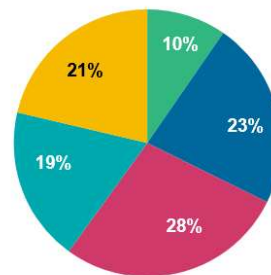
## Figures ...

Internet Users – 1995 → 2014...  
<1% to 39% Population Penetration Globally

**1995**  
35MM+ Internet Users  
0.6% Population Penetration



**2014**  
2.8B Internet Users  
39% Population Penetration



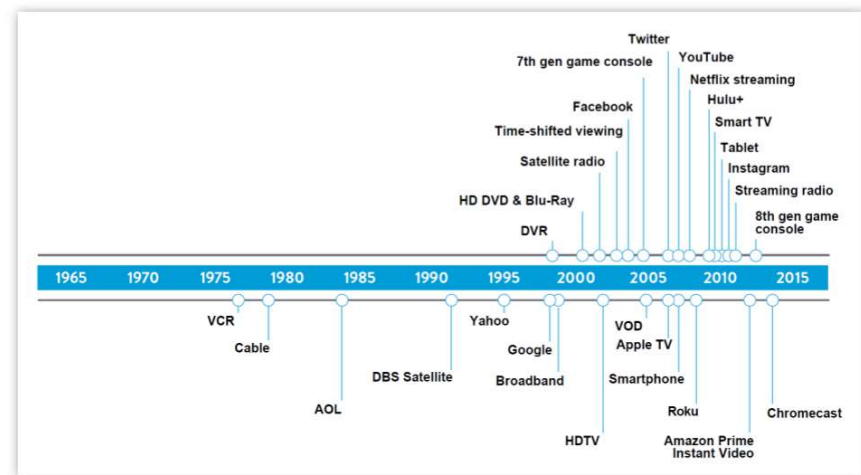
■ USA ■ China ■ Asia (ex. China) ■ Europe ■ Rest of World

@KPCB Source: Euromonitor, ITU, US Census

4

User Control of Content Up Significantly – 1995 → 2015

Evolution of Content Discovery, 1975 – 2015, per Nielsen



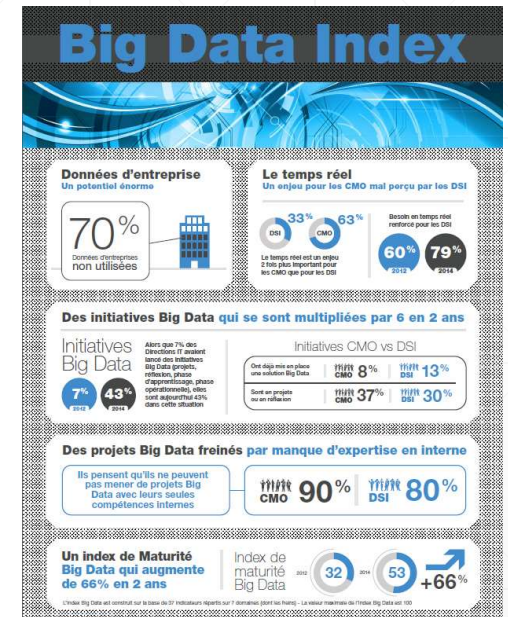
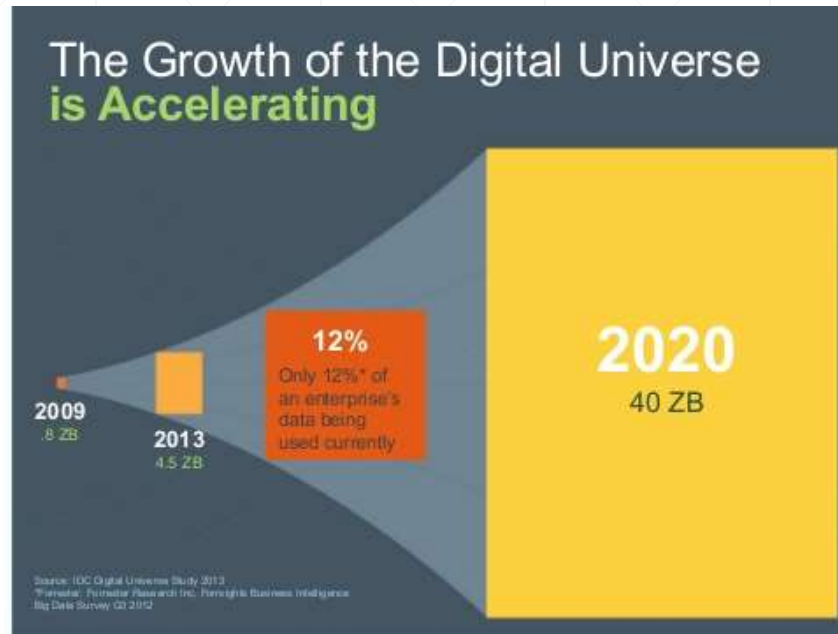
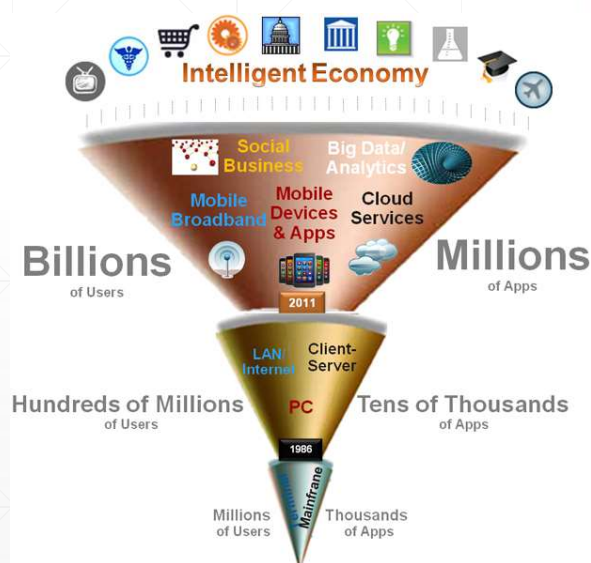
@KPCB Source: Nielsen, 12/14

7

Illustrations: KPCB. Internet Trends Report by Mary Meeker. May 2015.

# Today's Digital Economy = Data Economy

## More figures ...



<https://idc-community.com/energy/smart-grid/big-data-and-analytics-for-the-smart-grid>

<http://www.journaldunet.com/solutions/dsi/big-data-barometre-idc-emc-132014.shtml>

<http://fr.slideshare.net/MapRTTechnologies/webinar-fast-and-furious-from-pocto-enterprise-big-data-stack-20140424>



# Big Data & AI | Supply-side considerations

## ... it's a complex world ...

### BIG DATA & AI LANDSCAPE

#### Discussion:

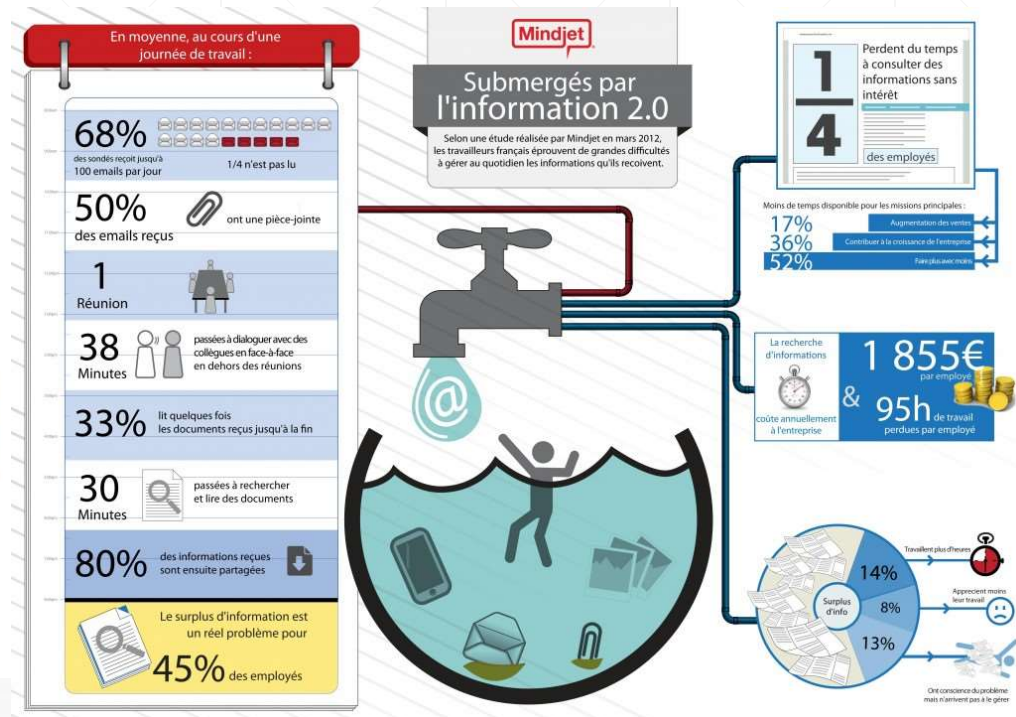
- A large variety of frameworks, architectures, technologies, applications, ..
- Frameworks
- Storage
- Databases
- Data Science
- Analytics & AI
- Industry-specific solutions
- Open vs proprietary standards, ...

# Big Data Fundamentals

## Technologies and environments

- Hadoop
  - NoSQL
  - NewSQL
  - Streaming
  - In-Memory processing
  - Parallel computing
  - ...
- Virtualization
  - On premise and in-the-cloud
  - Integration of several types of data sources
  - Real-time ... or not
  - Machine-to-machine versus machine-to-human orientation
  - ...

# Big Data & AI Information overload



Illustrations: <http://www.pavillon-orange.org/blog/?p=2530>

## Value (economics)

From Wikipedia, the free encyclopedia

*Not to be confused with [Market value](#).*

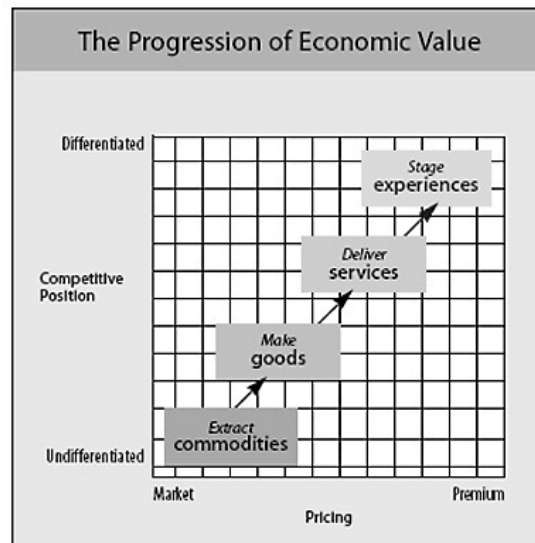
**Economic value** is a measure of the benefit provided by a [good](#) or [service](#) to an economic [agent](#). It is generally measured relative to units of [currency](#), and the interpretation is therefore "what is the maximum amount of money a specific actor is willing and able to pay for the good or service"?

Note that economic value is *not* the same as market price, nor is economic value the same thing as [market value](#). If a consumer is willing to buy a good, it implies that the customer places a higher value on the good than the market price. The difference between the value to the consumer and the market price is called "consumer surplus". It is easy to see situations where the actual value is considerably larger than the market price: purchase of [drinking water](#) is one example.

The economic value of a good or service has puzzled economists since the beginning of the discipline. First, economists tried to estimate the value of a good to an individual alone, and extend that definition to goods which can be exchanged.

From this analysis came the concepts [value in use](#) and [value in exchange](#).

[https://en.wikipedia.org/wiki/Value\\_\(economics\)](https://en.wikipedia.org/wiki/Value_(economics))



*Welcome to the Experience Economy.* By B. Joseph Pine II, James H. Gilmore.  
Published in the July–August 1998 HBR Issue  
<https://hbr.org/1998/07/welcome-to-the-experience-economy>

## Reminder: Value

“Value” measures the benefit that an economic agent (individual or firm) derives from a good or from a services. While a subjective concept in general, in economics, value is measured in currency units (monetary value).

In economic terms, one may also refer to the concept of “utility” which is the satisfaction an economic agent derives from the consumption of a good or service.

# The Information Society

## One paradox and one bottleneck

### RARITY => VALUE

- On the one hand, abundance of something creates a scarcity in other areas
- Whenever something is scarce, then there is value in providing the rare item

### EFFICIENCY/PRODUCTIVITY

- On the other hand, the question arises on how to organize efficiently in order to cater – in the most effective way - for the needs to be fulfilled through scarce goods or services

# Data from an economic perspective

## Two metaphors

CURRENCY

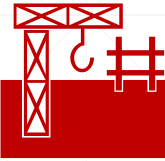


OIL



Source: [https://fr.wikipedia.org/wiki/Puits\\_de\\_p%C3%A9trole#/media/Fichier:Puits\\_de\\_p%C3%A9trole.jpg](https://fr.wikipedia.org/wiki/Puits_de_p%C3%A9trole#/media/Fichier:Puits_de_p%C3%A9trole.jpg)



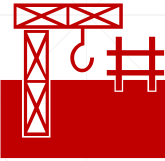


## Class Assignment

Drawing on today's reading material and the discussion so far, please comment on the following:

- a) The possible combinations between the various forms of **decision-making** described in the reading material,
- b) The role of data and AI

... in the following business settings:



# Class Assignment Business Settings

1. A **health professional** providing diagnostics and prescriptions
2. An **insurance company** providing health insurance services
3. A **financial services company** providing credit scoring services to other companies
4. A **recommendation engine** providing « cross-sell » and « up-sell » information
5. A **targeted advertising service** providing personalized ads across multiple devices and channels
6. A **retail merchant** with both online and physical presence
7. A « **driverless car** » **service** in the context of driving in heavy traffic
8. An **industrial manufacturing company** providing maintenance services

# Where do we stand with Big Data? (as we speak) 2013 vs 2020 (7 years' time lapse)

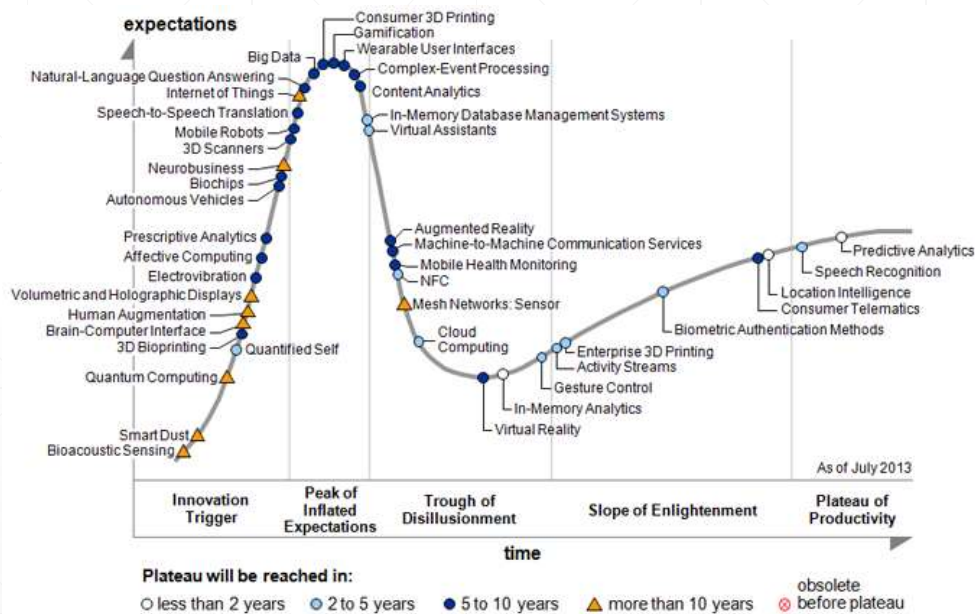


Chart: Gartner's 2013 Hype Cycle for Emerging Technologies  
<http://www.gartner.com/newsroom/id/2575515>

## Hype Cycle for Emerging Technologies, 2020



[gartner.com/SmarterWithGartner](https://www.gartner.com/SmarterWithGartner)

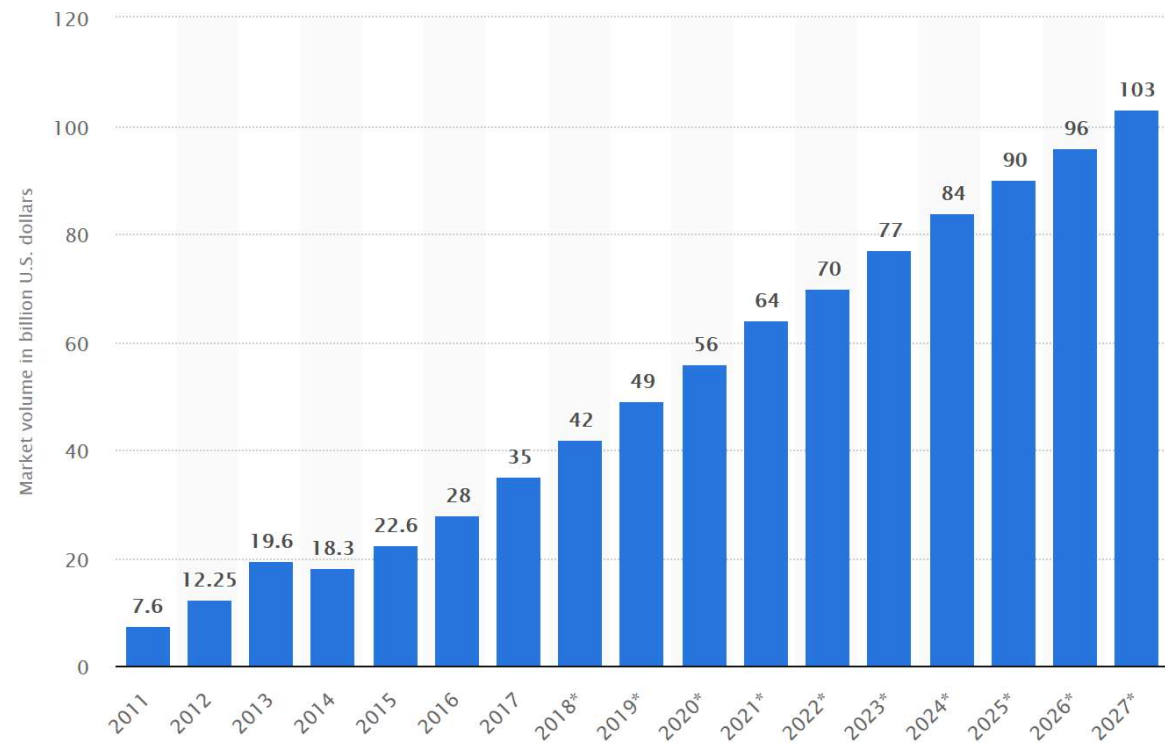
Source: Gartner  
© 2020 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and Hype Cycle are registered trademarks of Gartner, Inc. and its affiliates in the U.S.

**Gartner.**

<https://www.blogdumoderateur.com/gartner-hype-cycle-emerging-technologies-2020/>

# Big Data & AI | Demand-side considerations

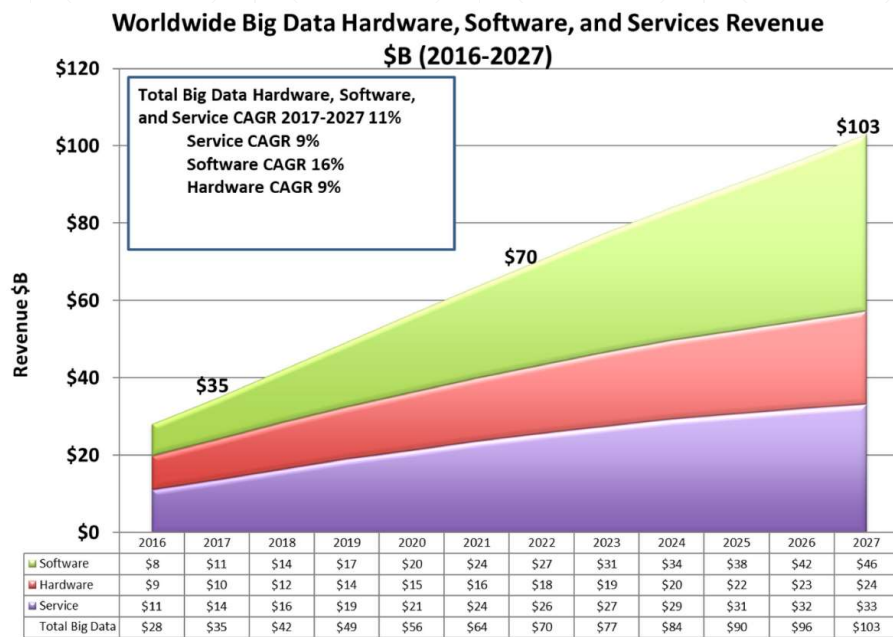
## Aggregated demand



Source: Statista. <https://www.statista.com/statistics/254266/global-big-data-market-forecast/>

# Data | Supply-side considerations

## Global Market Forecast



- The chart on the left (from Wikibon's 2018 Report) shows the projected growth of the market for Big Data
- Insights (consistent with previous editions):
  - Among the **main drivers for growth**, the report quotes increased cooperation/concentration between vendors
  - In terms of **adoption barriers**, the report emphasizes the lack of practices to integrate Big Data analytics and AI into existing processes

Chart: Wikibon. <https://wikibon.com/wikibons-2018-big-data-analytics-trends-forecast/>

# Big Data & AI | Supply-side considerations

## Databases

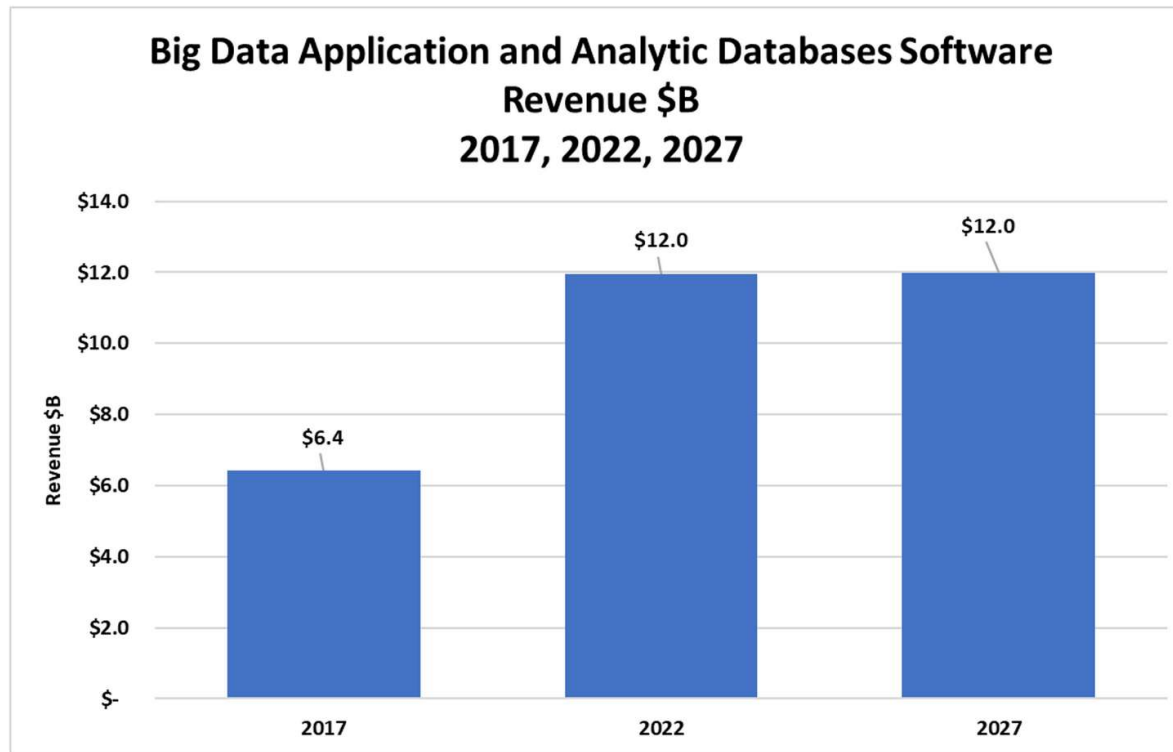


Chart: Wikibon. <https://wikibon.com/wikibons-2018-big-data-analytics-trends-forecast/>



# Big Data & AI | Supply-side considerations

## Software Analytics

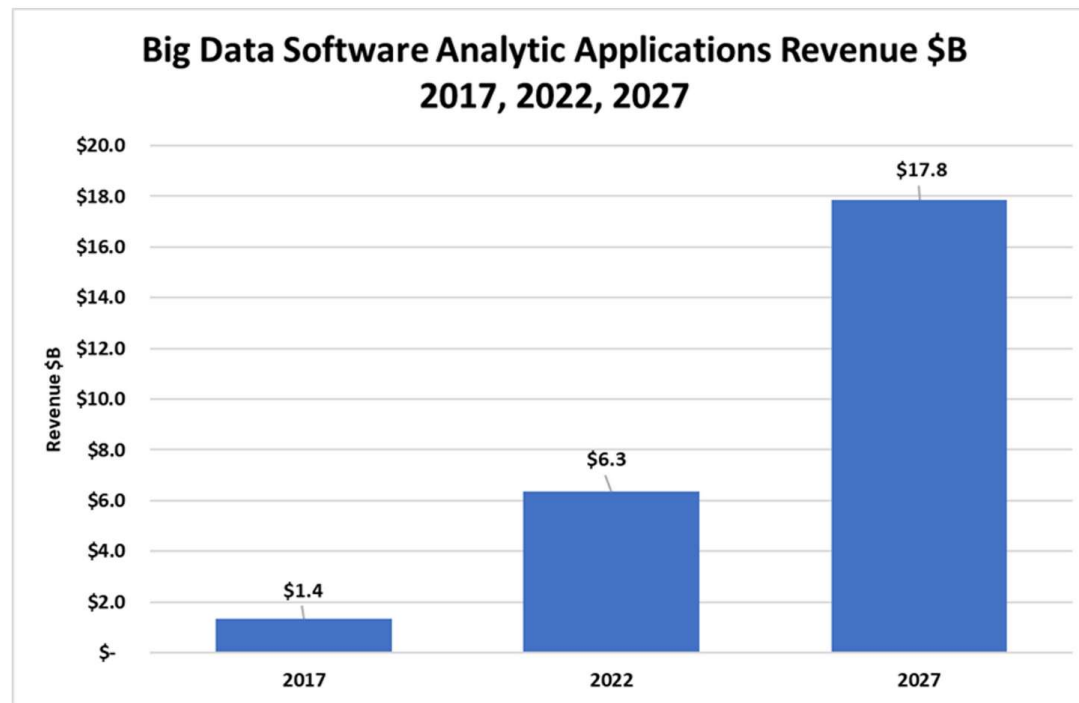


Figure 4: Big Data Analytic Applications Software Revenue \$B – 2017, 2022, 2027

Chart: Wikibon. <https://wikibon.com/wikibons-2018-big-data-analytics-trends-forecast/>

# Big Data & AI | Supply-side considerations

## Data Science

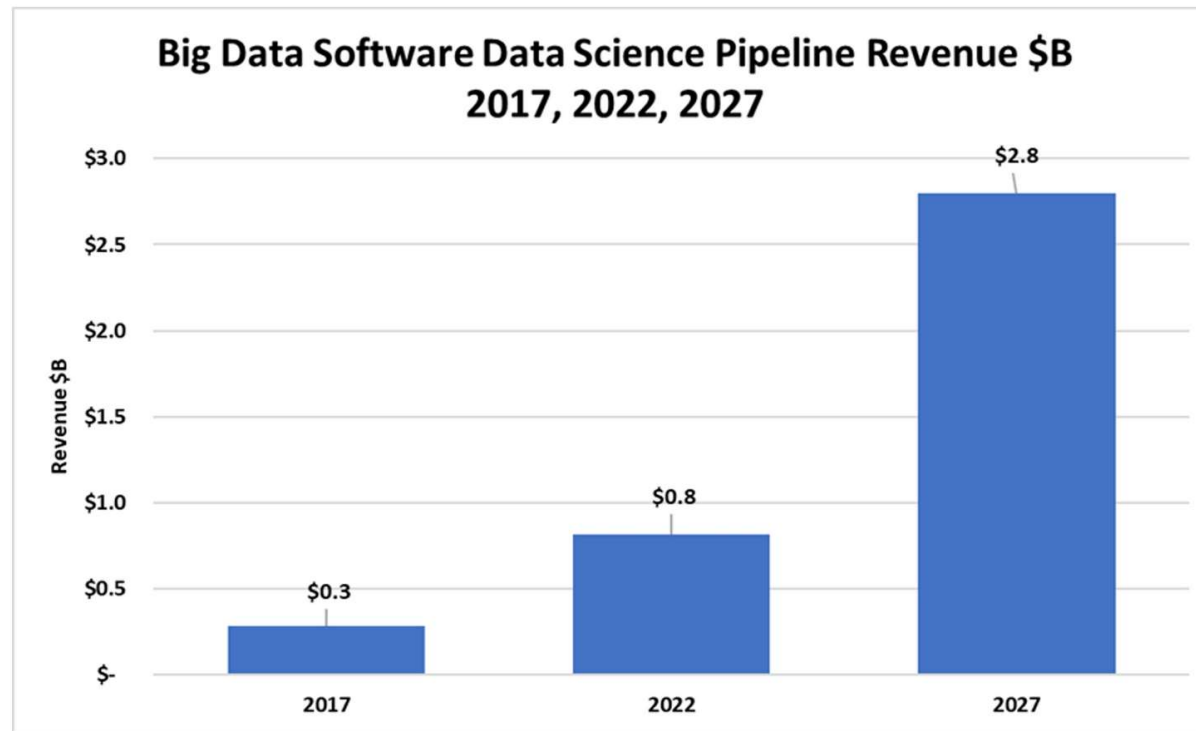


Figure 5: Big Data Data Science Pipelines Software Revenue \$B – 2017, 2022, 2027

Chart: Wikibon. <https://wikibon.com/wikibons-2018-big-data-analytics-trends-forecast/>

# Big Data & AI | Supply-side considerations

## Professional Services

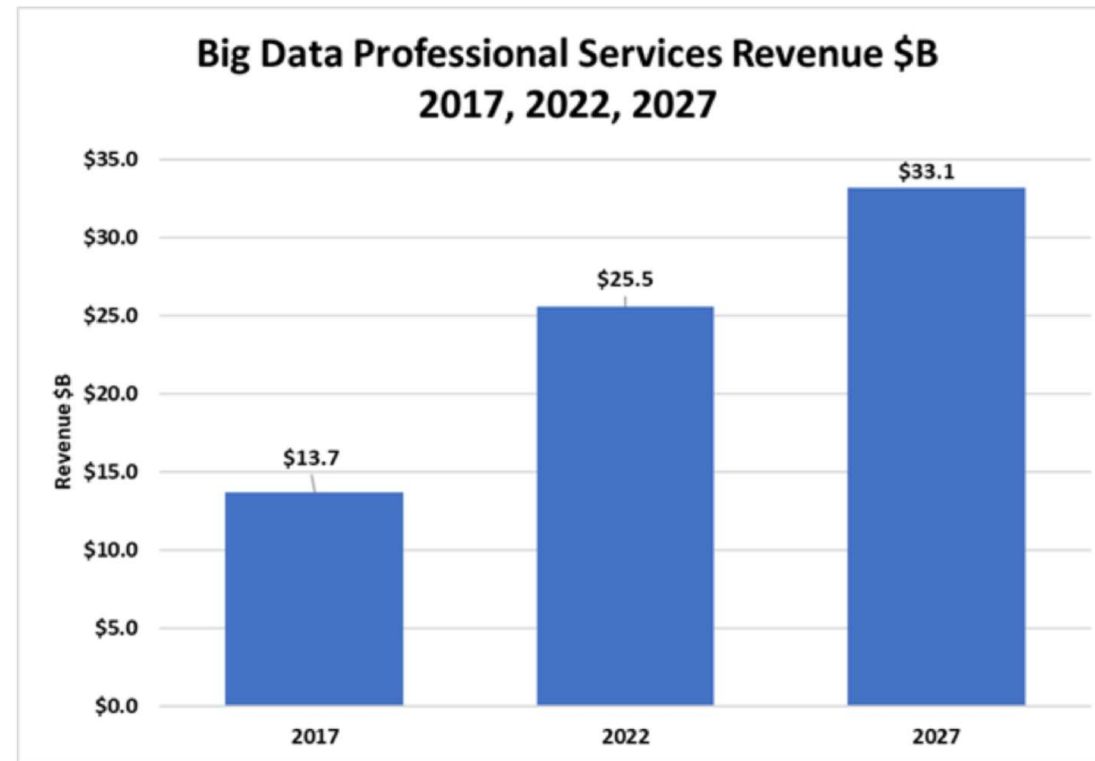
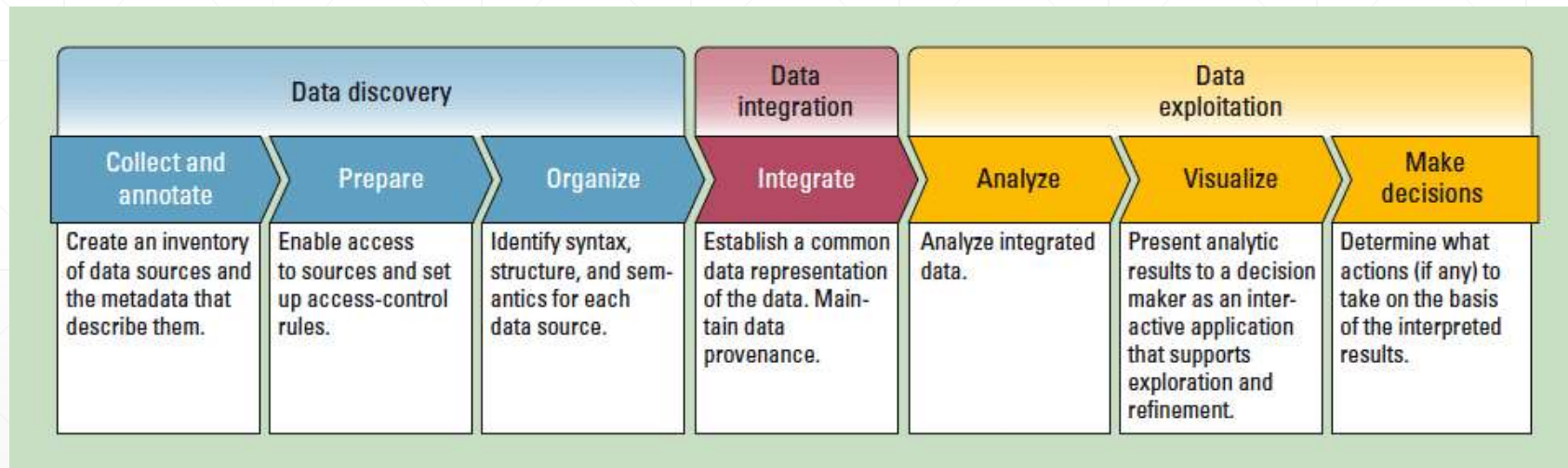


Chart: Wikibon. <https://wikibon.com/wikibons-2018-big-data-analytics-trends-forecast/>

# Big Data & AI| Extracting value

## What is the proposed value chain?



H. Gilbert Miller, Peter Mork, "From Data to Decisions: A Value Chain for Big Data", *IT Professional*, vol.15, no. 1, pp. 57-59, Jan.-Feb. 2013, doi:10.1109/MITP.2013.11<http://www.computer.org/csdl/mags/it/2013/01/mit2013010057-abs.html>

# Big Data & AI Discussion



## Value Chain

Value Creation
Data Generation
Data acquisition & discovery <ul style="list-style-type: none"><li>- Ability to store large volumes of data</li><li>- Ability to organize data from various sources</li></ul>
Data integration <ul style="list-style-type: none"><li>- Ability to establish a common view of disparate data</li></ul>
Data Exploitation <ul style="list-style-type: none"><li>- Ability to extract useful information from data</li><li>- For one own's purpose or for third parties</li></ul>

## Players

Types of Companies
Technology companies – HW SW
Consulting and Systems Integration
Internet Players (large, ex.: GAFA)
Internet players (small and medium)
Industry (Manufacturing)
Telecommunications
Retail (including e-commerce)
Media
Financial Services (Banks, Insurance, etc)
Government

# Big Data & AI | Demand-side Considerations

## How Companies Are Using Big Data & AI

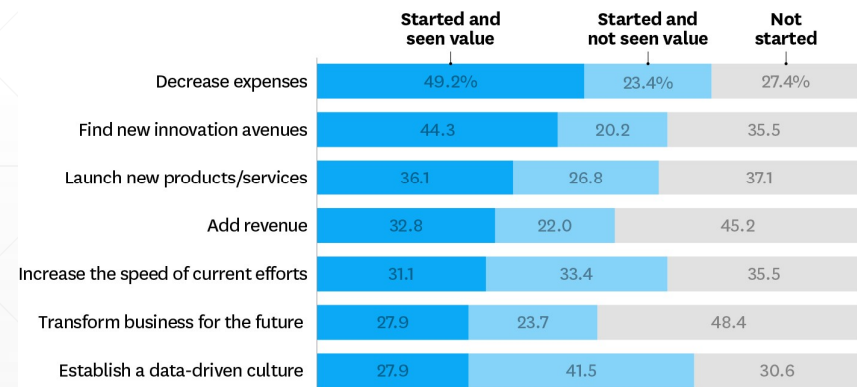
### By Business Use Case

- 360° View of the customer
- Fraud prevention
- Security intelligence
- Datawarehouse/Business Intelligence Applications
- Price Optimization
- Operational Efficiency
- Recommendation engines
- Social Media Analysis and Response (sentiment analysis)
- Preventive Maintenance and support

### By Business Objective

#### How Fortune 1000 Executives Report Using Big Data

The projects they've started, and where they're finding value.



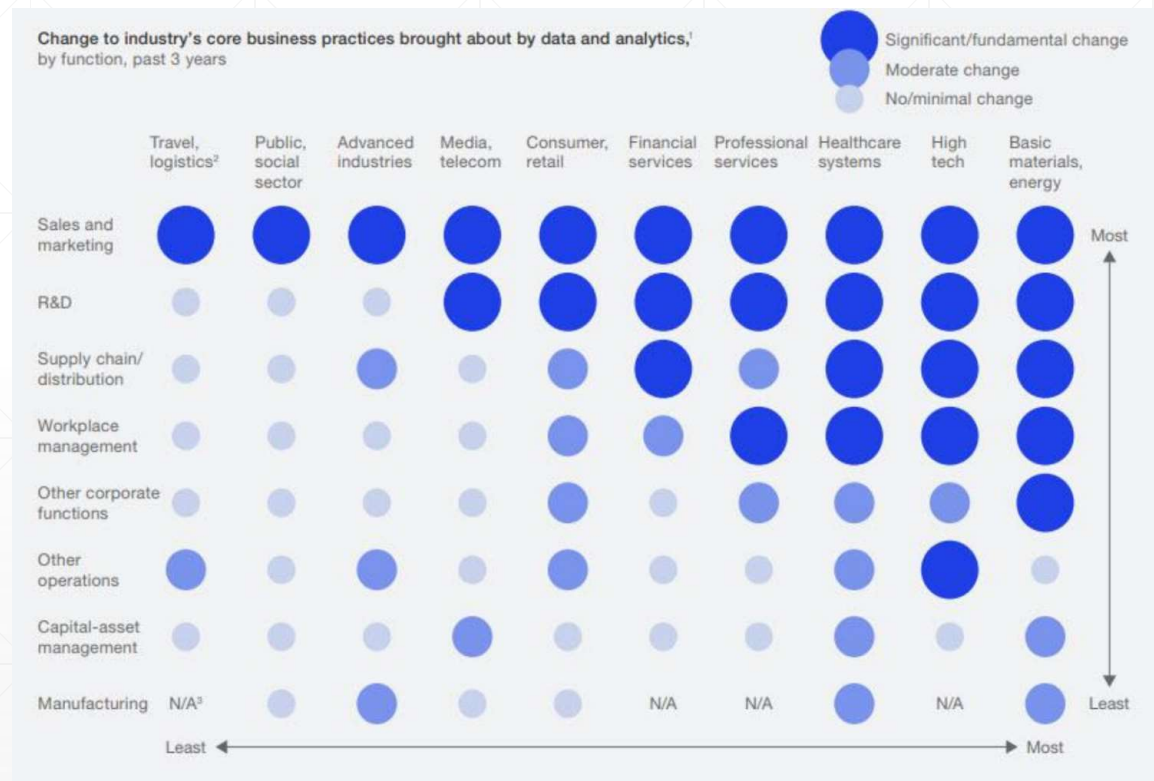
SOURCE: NEWVANTAGE PARTNERS BIG DATA EXECUTIVE SURVEY, 2017

© HBR.ORG

Source: Randy BEAN. *How Companies Say They're Using Big Data*. Harvard Business Review, April 28, 2017.

# Big Data & AI | Demand-side considerations

## Change - by industry and business function

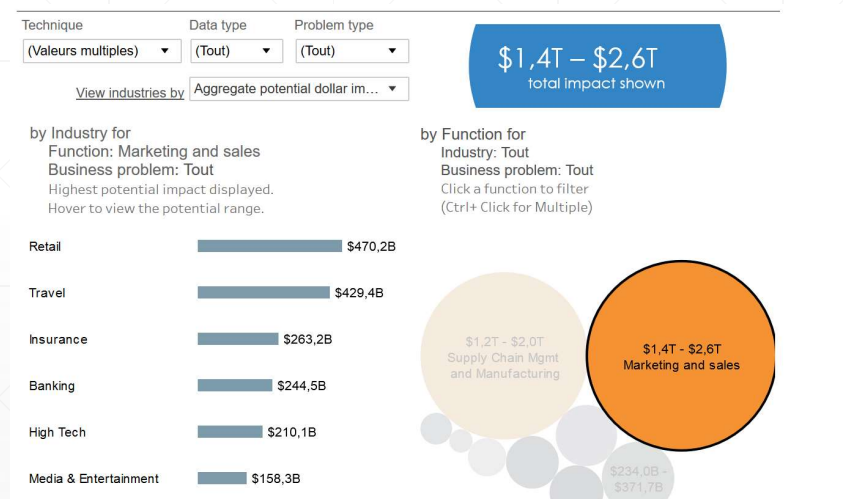


Source: McKinsey&Company. McKinsey Analytics. *Analytics comes of age*. January 2018



# Big Data & AI | Demand-side considerations

## AI & Analytics - Marketing & Sales



Source: McKinsey&Company. McKinsey Analytics.  
<https://www.mckinsey.com/featured-insights/artificial-intelligence/visualizing-the-uses-and-potential-impact-of-ai-and-other-analytics>

- Customer Service Management
- Next Product to Buy, Next Best Offer, Individualized offerings
- Pricing and Promotion
- Churn Reduction
- Channel Management
- Customer Acquisition
- Marketing Budget Allocation

# Big Data & AI

## Internal factors influencing adoption [1/3]

Intrinsic (technical) barriers to ML adoption	Description
<b>Labeling Data</b>	In supervised learning, data sets need to be enriched with outcome values. This activity is, at a time, extremely time-consuming, it requires massive investments in human labor, and may generate errors
<b>Access to large training data sets</b>	In ML, the accuracy of outcomes and the preciseness of results is largely dependent on the size of the training data sets that computer systems need to 'ingest.' Such large data sets are rarely available, generally speaking, and, for some specific cases (for instance, in medical applications) building such comprehensive training data ensembles may turn out to be gruesomely difficult.

# Big Data & AI

## Internal factors influencing adoption [2/3]

### Intrinsic (technical) barriers to ML adoption

### Description

#### Explainability

AI, and, in particular ML techniques appear, very often, to non-specialist users to be a “black-box”-type of a business-support system. In many circumstances, particularly when the expected output informs decision-making, this opaqueness is prone to generate reluctance in ML adoption.

#### Generalizability

ML models are, traditionally, very specialized. Given their relatively high initial cost, this over-specialization is clearly a barrier to their wider adoption by any organization. From an investment perspective, technical solutions need to be found in order to increase the applicability of ML models to two, or more, types of problems to be solved.

# Big Data & AI

## Internal factors influencing adoption [3/3]

Intrinsic (technical) barriers to ML adoption	Description
<b>Bias(*) in data and algorithms</b>	By far, the most impeding hurdle to ML adoption, bias may be rooted both in data and in the computer model. In both cases, it generates repeatable errors that may lead to erroneous, unfair, discriminatory, potentially dangerous, or illegal conclusions and/or actions.

(\*) Bias, in the context of AI, is related to the likelihood of, systematically, producing erroneous results, or reaching wrong conclusions.

For more information on bias and ethics in AI, please see:

<https://www.institutmontaigne.org/publications/algorithmes-controle-des-biais-svp>

# In short

## The role of data

### Data as in “oil”

- Input
- Value (supply) chain:
  - Extraction
  - Refinement
  - Distribution/Delivery

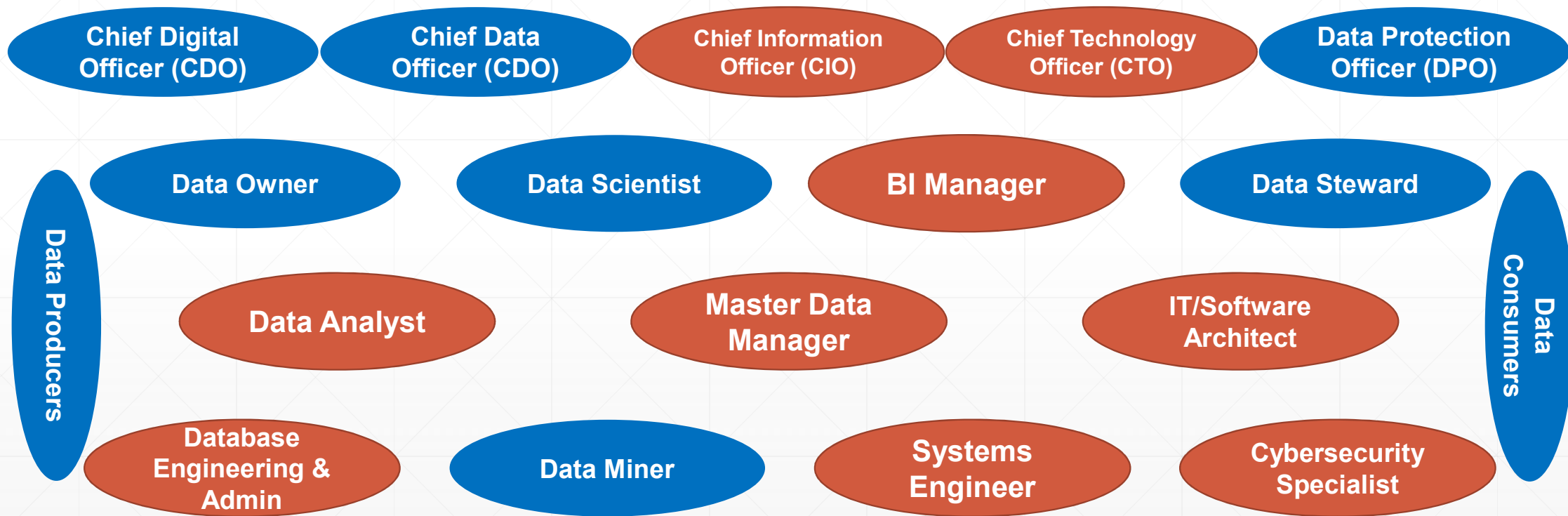
### Data as in “currency”

- Data as sellable good (e.g. data-brokers)
- Data as information for decision and action (i.e. optimize existing operations, business processes and decision-making)
- Data as strategic asset (e.g. leverage data to transform the business, ex: become a digital platform or to conquer new markets)

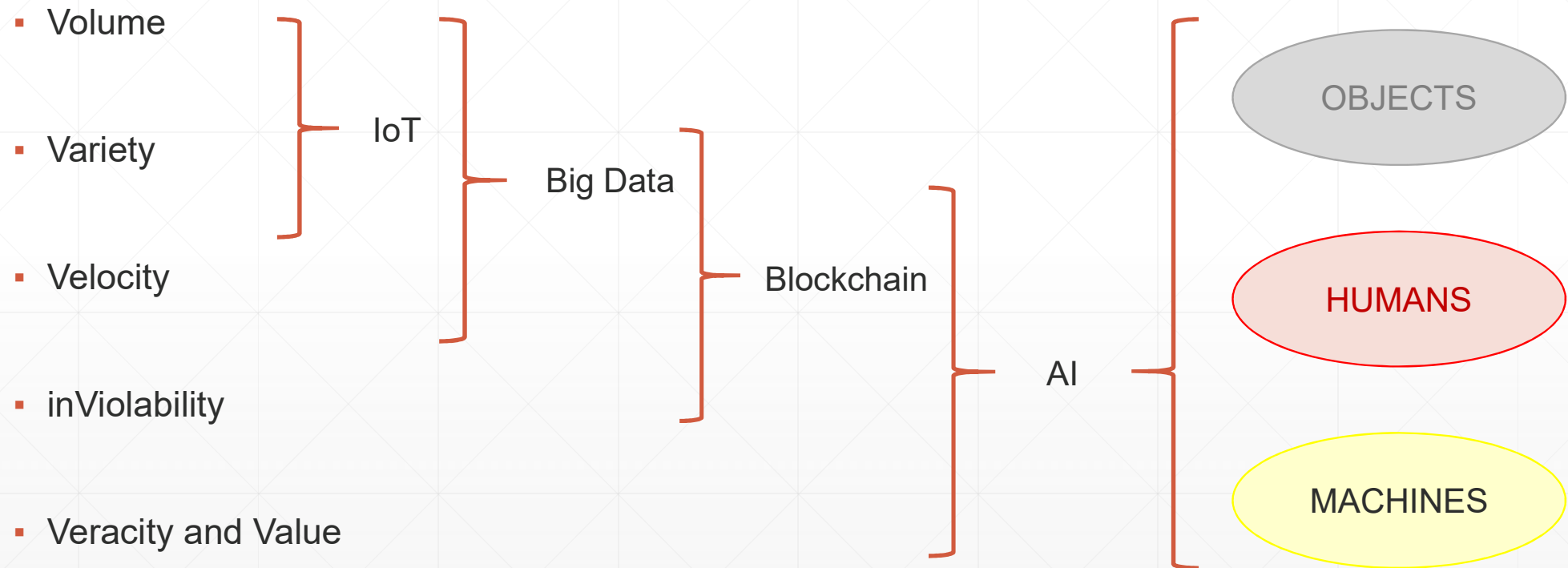
**Data => Information => Knowledge**

# Data-Driven Organizations

## Existing vs New Roles



# IoT, Blockchain, Big Data & AI Combinatorial Effects - I





# IoT, Blockchain, Big Data & AI Combinatorial Effects - II

- Abundance of Data
- ... creates the need to:
  - Build trust in a decentralized and distributed global infrastructure
  - Design and implement the technical environments capable of handling volume, variety, velocity, etc...
  - Transform data into relevant and trustworthy information and knowledge

IoT

Block  
chain

Big  
Data

AI

FIGURE 22

Top 20 job roles in increasing and decreasing demand across industries

➤ Increasing demand

1	Data Analysts and Scientists
2	AI and Machine Learning Specialists
3	Big Data Specialists
4	Digital Marketing and Strategy Specialists
5	Process Automation Specialists
6	Business Development Professionals
7	Digital Transformation Specialists
8	Information Security Analysts
9	Software and Applications Developers
10	Internet of Things Specialists
11	Project Managers
12	Business Services and Administration Managers
13	Database and Network Professionals
14	Robotics Engineers
15	Strategic Advisors
16	Management and Organization Analysts
17	FinTech Engineers
18	Mechanics and Machinery Repairers
19	Organizational Development Specialists
20	Risk Management Specialists

➤ Decreasing demand

1	Data Entry Clerks
2	Administrative and Executive Secretaries
3	Accounting, Bookkeeping and Payroll Clerks
4	Accountants and Auditors
5	Assembly and Factory Workers
6	Business Services and Administration Managers
7	Client Information and Customer Service Workers
8	General and Operations Managers
9	Mechanics and Machinery Repairers
10	Material-Recording and Stock-Keeping Clerks
11	Financial Analysts
12	Postal Service Clerks
13	Sales Rep., Wholesale and Manuf., Tech. and Sci.Products
14	Relationship Managers
15	Bank Tellers and Related Clerks
16	Door-To-Door Sales, News and Street Vendors
17	Electronics and Telecoms Installers and Repairers
18	Human Resources Specialists
19	Training and Development Specialists
20	Construction Laborers

Source

Future of Jobs Survey 2020, World Economic Forum.

**Source:** WEF. [http://www3.weforum.org/docs/WEF\\_Future\\_of\\_Jobs\\_2020.pdf](http://www3.weforum.org/docs/WEF_Future_of_Jobs_2020.pdf)

# The class/reference example

## IoT-Blockchain-Big Data-AI

- **IoT** is the technology enabling objects to be connected to a network, either existing, or new
  - In our case, the reader of reference and student IDs and the portable ID for each “item”
- **Blockchain** is the technology allowing trust to be embedded in the system
  - In our case, the rules that make sure that all references are returned at the end, and, that students may exchange references between them without going through a central point (professor)
- **Big Data** is the technology that allows variety of references (books, articles, videos, pictures, audio recordings), volume (Zettabytes) and velocity (real time)
  - In our case, making information about multiple types of references usage available in real time, and the possibility to create a central, digital, catalogue
- **Artificial Intelligence** is the set of rules that optimizes the usage of references by students for improved learning experience
  - In our case, providing this year’s class with information about prior usage (historical), and recommendation on the best reference combinations depending on the final subject presentations

# IoT, Blockchain, Big Data & AI - Wrap-Up

## Adoption considerations

- Combination of several technologies
- Adoption discussion:
  - Standards (ex: MS-DOS)
  - Installed base (ex: QWERTY keyboard)
  - Convergence between different types of knowledge
    - (ex: computer: binary arithmetic, symbolic logic, programming)
  - Co-existence of what initially was considered to be rival technologies (ex: DC and AC)

# Big Data & AI

## Class 3 Summary



- Big Data & AI is a powerful combination of technologies
  - Work structure inherited from the past is questioned, as we are in the presence of Man-Machine interactions of a new type. Decision making processes are increasingly spread across people, processes and technology
- There are internal factors influencing adoption, the most important being: bias, explainability, and generalizability

***“The most important general-purpose technology of our era is artificial intelligence (AI), particularly machine learning (ML).”***

(Brynjolfsson & McAfee, 2017)

# Thank You!

---

Valeriu Petrulian