Explosive Advances in Computer Science:
with a focus on Data Science, Cybersecurity, and Cloud
—

Dr. Natesan Venkateswaran
Senior Engineering Manager, IBM Systems
IBM Master Inventor, Member – IBM Academy of Technology

DISCLAIMER

Some material in this deck is publicly available information, used solely for informational purposes

In most cases, the relevant sources are cited

Any missing source is inadvertent, and not intentional

Evolution of Computing

https://www.gregverdino.com/everything-is-exponential/

Evolution of Computer Power/Cost

https://frc.ri.cmu.edu/~hpm/talks/revo.slides/power.aug.curve/power.aug.html
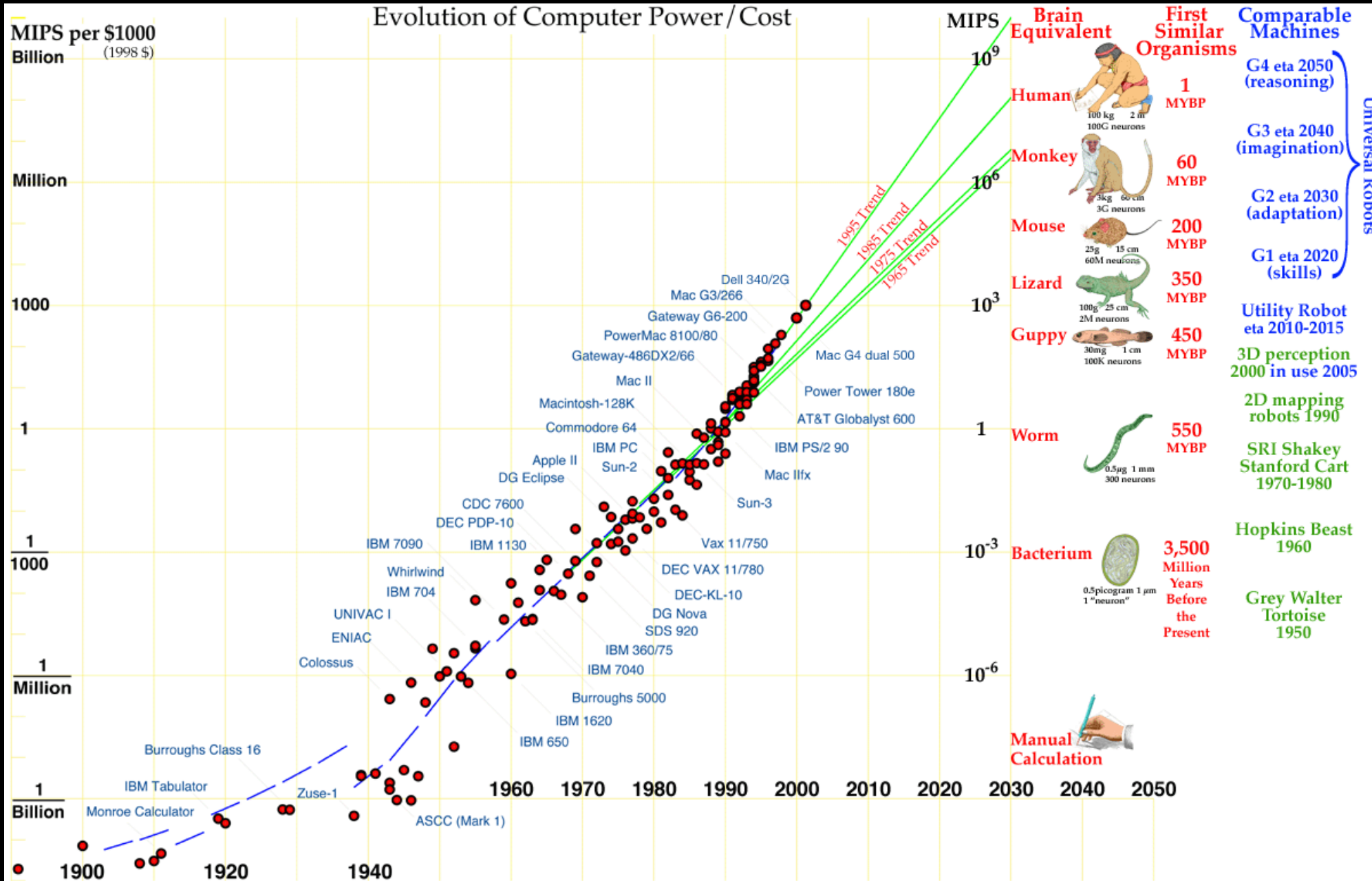
Figure 2. Technology costs are plummeting (and the reach is increasing)

Compute
- 1961: $1.1T
- 2015: $0.08
($/GFLOPS, 1960–2020)

Storage
- 1964: $3.5B
- 2015: $27
($/TB, 1960–2020)

Connectivity
- 1998: $1200
- 2015: $0.63
- Wireline, US only
($/Mbps, 1995–2020)
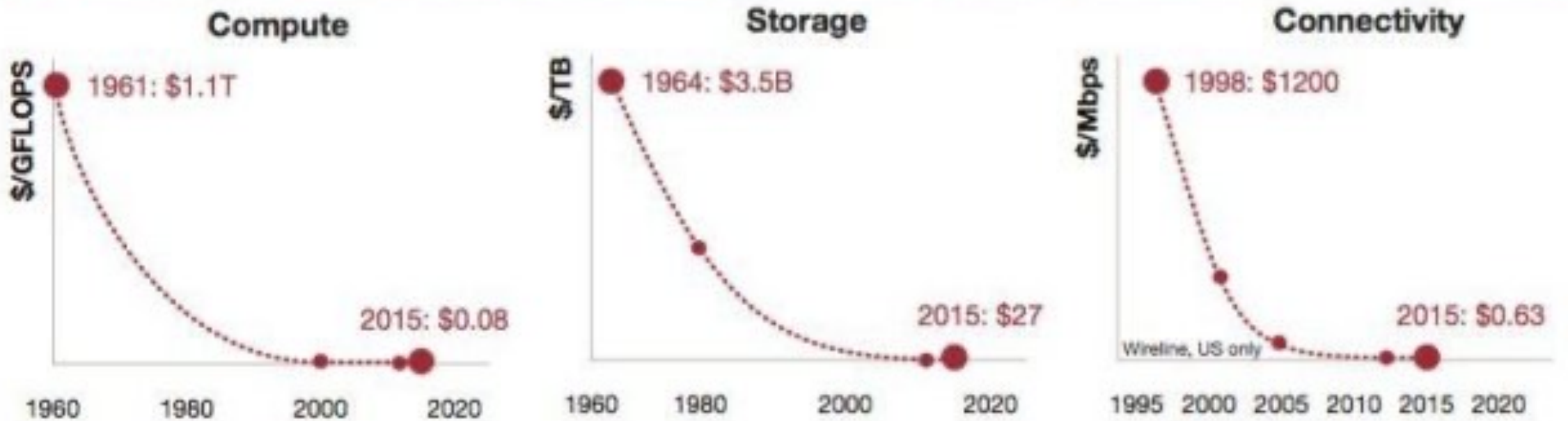
Source: PwC/Strategy& analysis; Michael Driscoll/Metamarkets

https://whatsthebigdata.com/2017/02/24/8-new-technologies-driven-by-plummeting-costs-of-computing-storage-and-networking/

The three eras of computing

https://blogs.opentext.com/future-information-rise-cognitive/
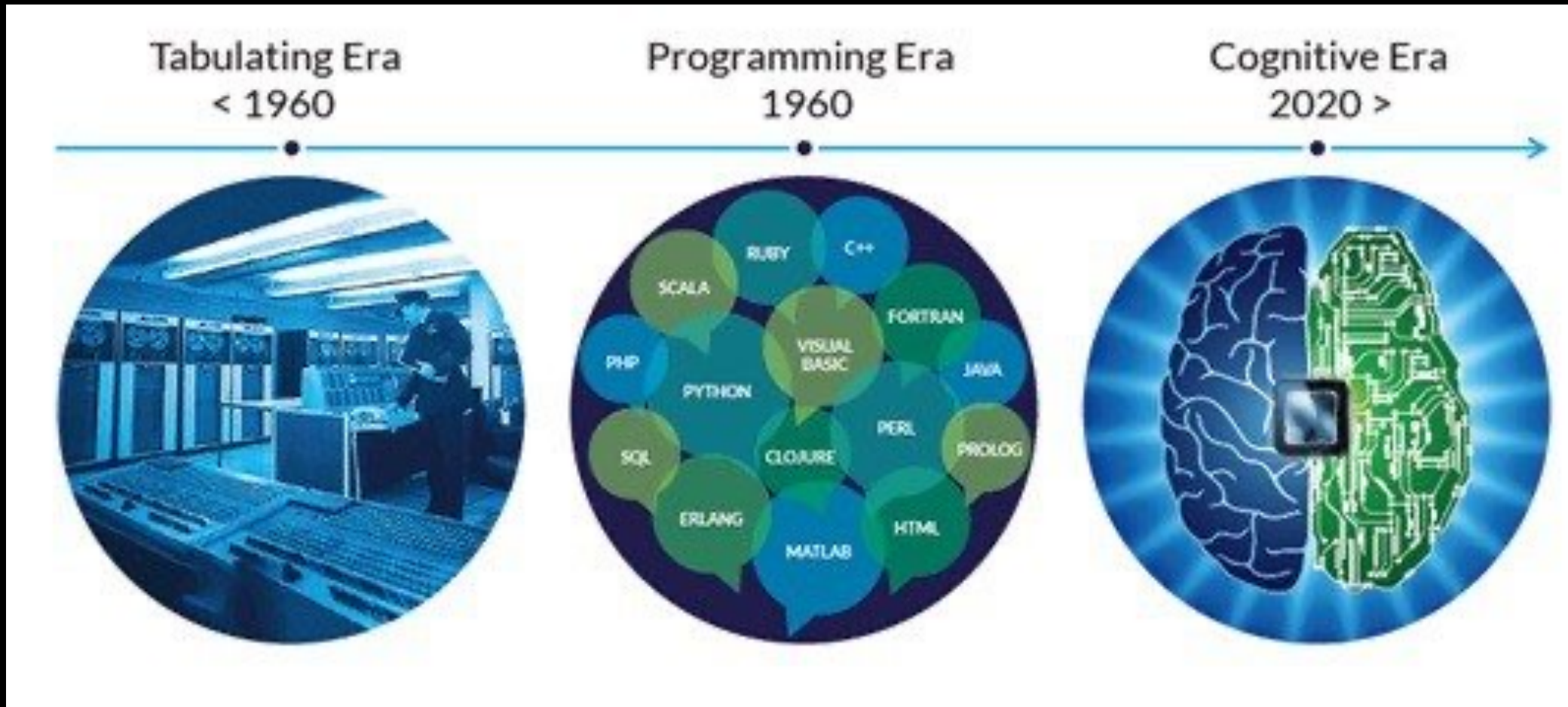
Data Driven Era

The Journey to Intelligence

The development of artificial intelligence: Rapid transition thanks to Mobile, Cloud and Big Data

Artificial Intelligence

Big Data

Mainframe computing **Client server** **Internet** **Mobile-Cloud**

2020s

1950   1980   2000   2010

Programmable era

Intelligence era
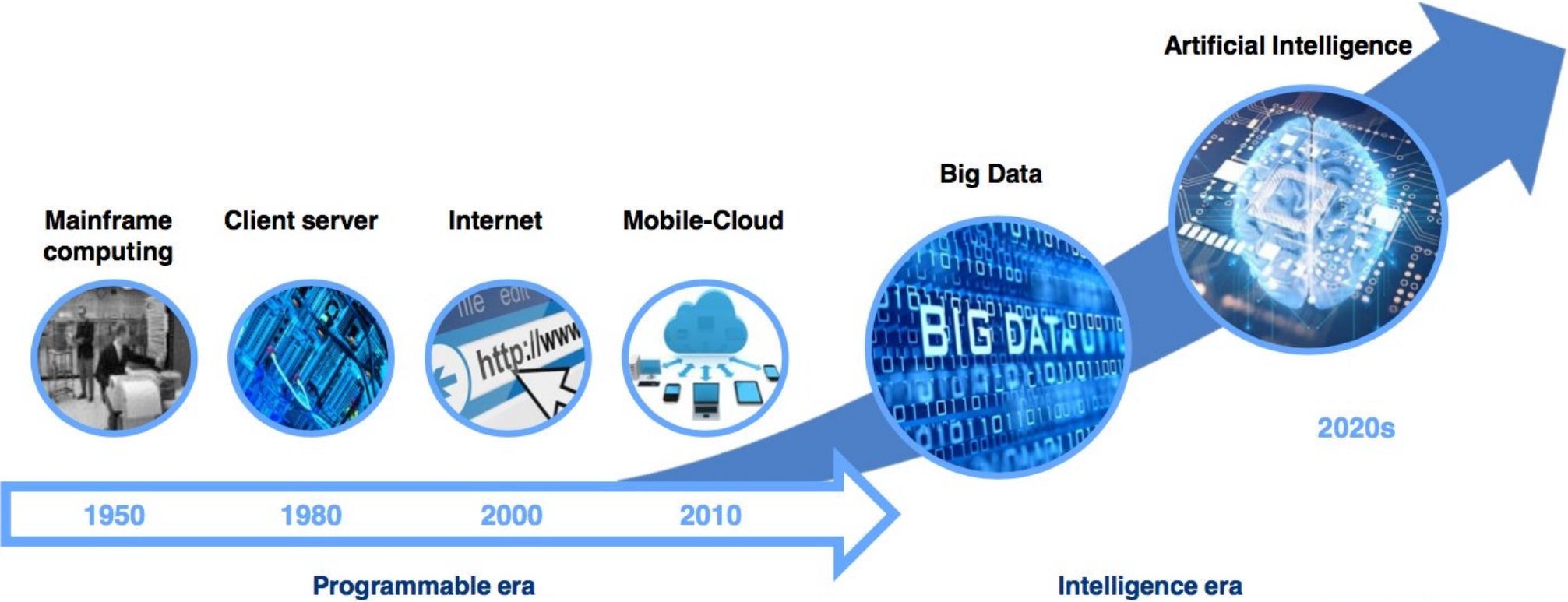
source allianz via @mikequindazzi
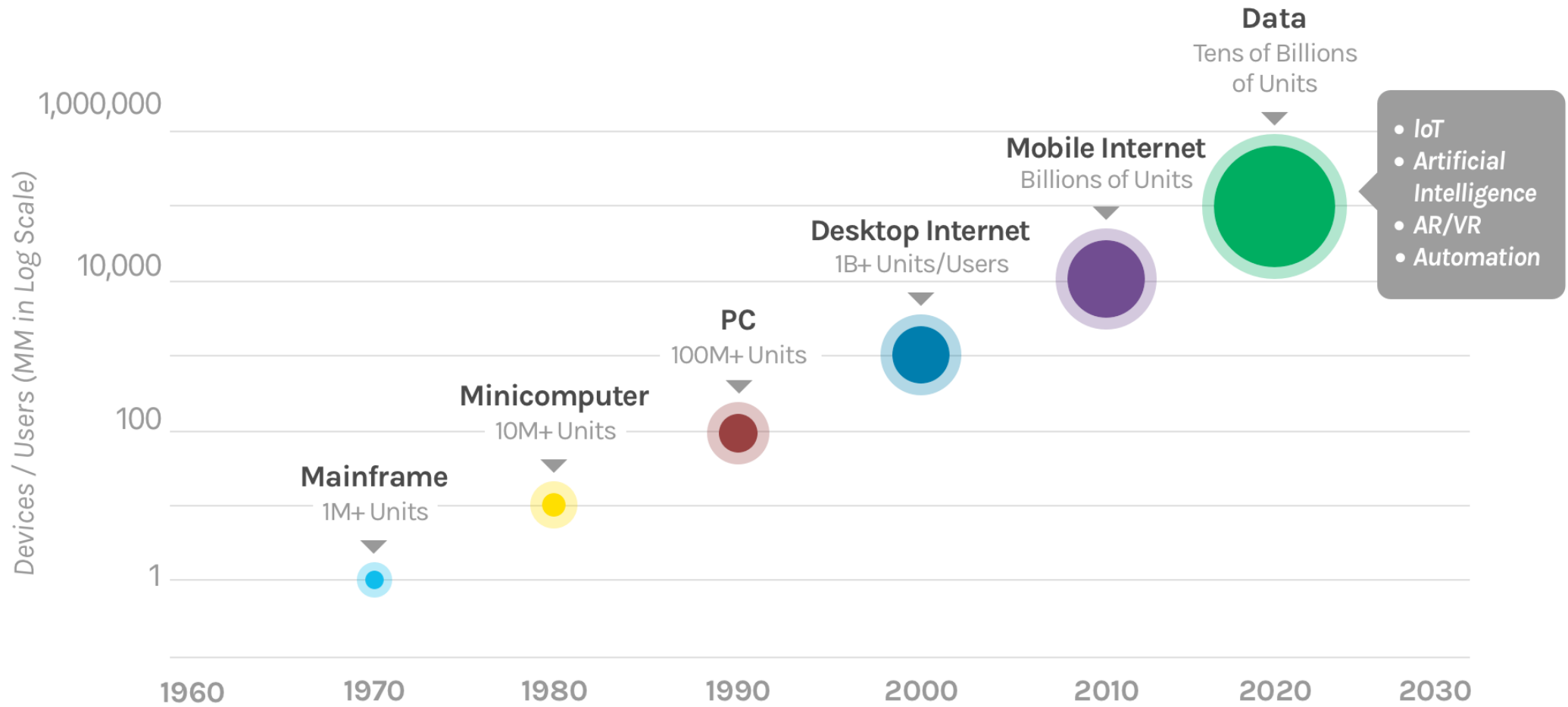
https://twitter.com/mikequindazzi/status/945248511976763393

https://www.morganstanley.com/ideas/data-decade/
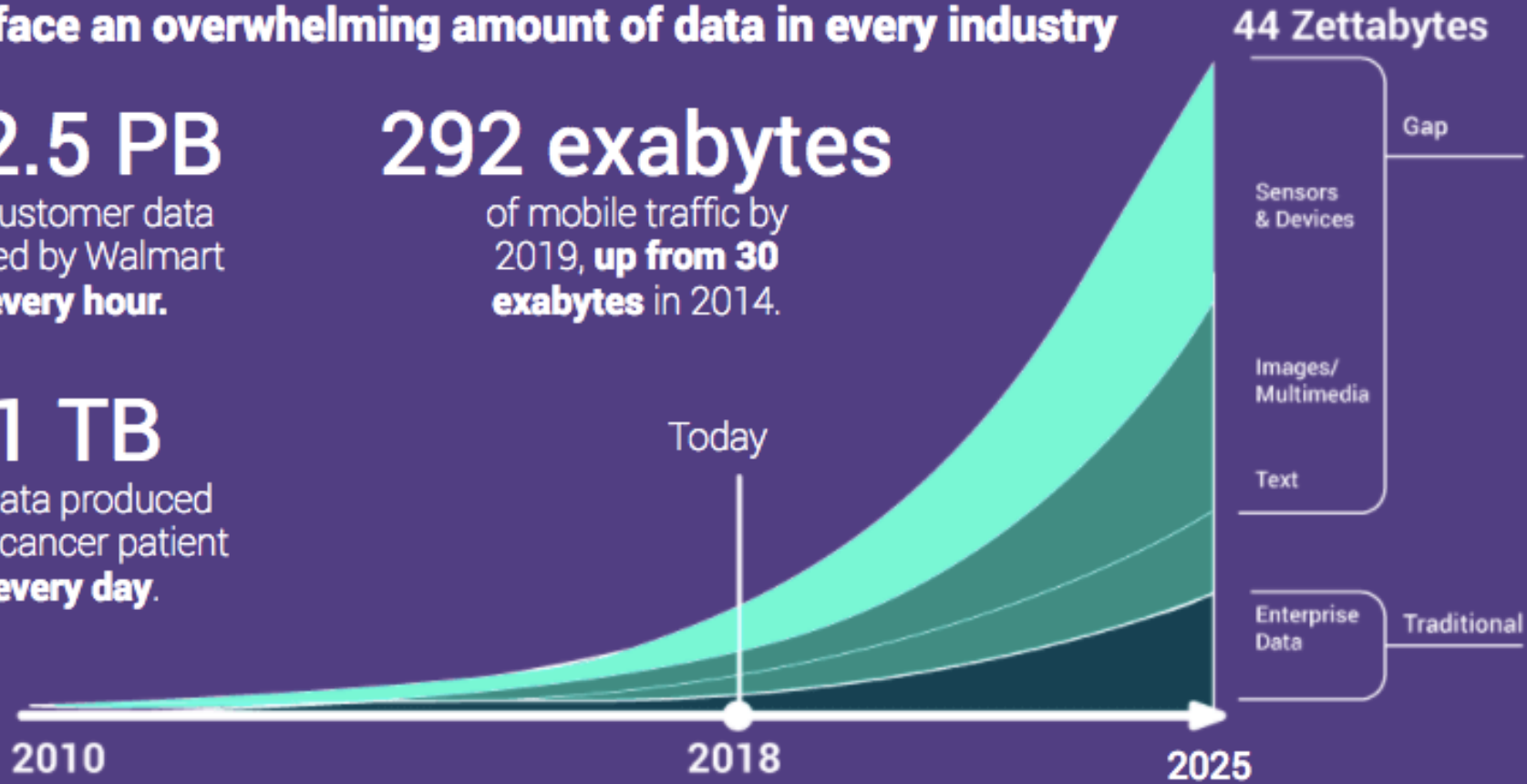
https://financesonline.com/how-much-data-is-created-every-day/

Trends 2022

Security

Cloud

Data Science

Top Strategic Technology Trends for 2022

Data Fabric
Composable Applications
Distributed Enterprise
Cybersecurity Mesh
Decision Intelligence
Total Experience
Privacy-Enhancing Computation
Hyperautomation
Autonomic Systems
Cloud-Native Platforms
AI Engineering
Generative AI
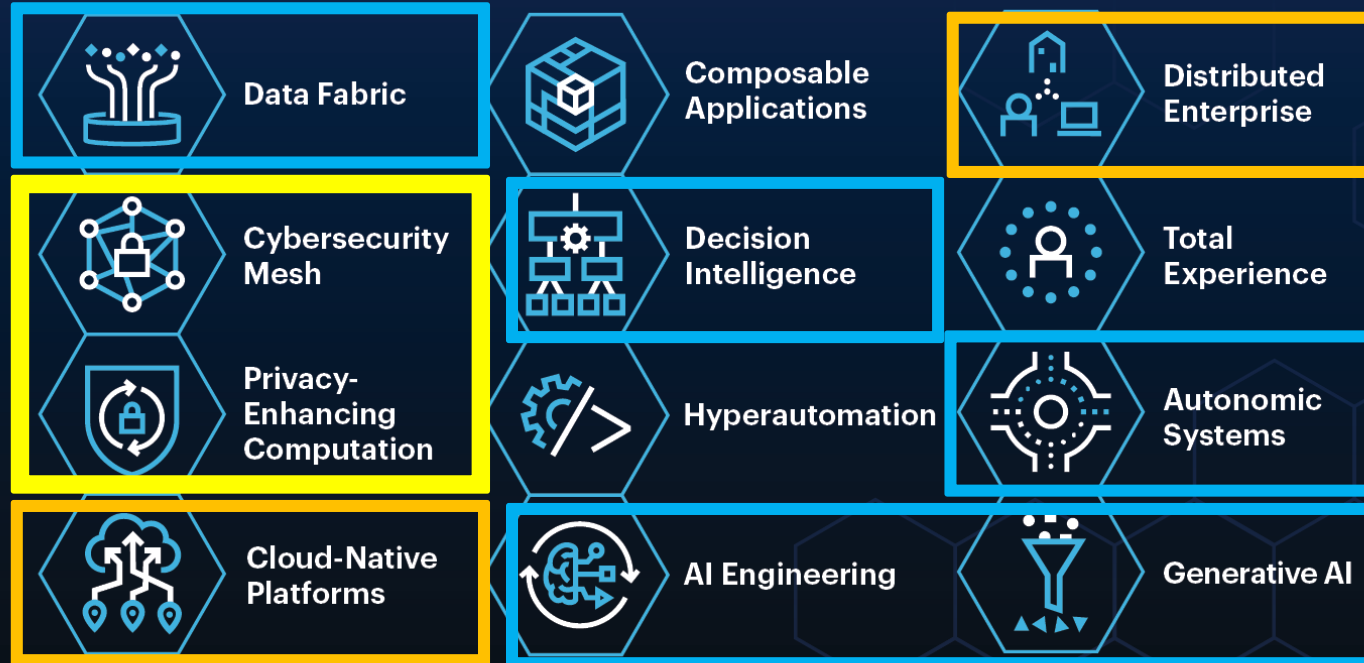
© 2021 Gartner, Inc. and/or its affiliates. All rights reserved. 1397600

Gartner

https://www.gartner.com/en/information-technology/insights/top-technology-trends

14

What is Data Science?

Data is inherent in every aspect of our lives

- Smartphones, IoT, Internet, Sensors, Alexa, …

Data is like water in ocean

- There is so much of it, but we cannot benefit from data until we process it

Data Science is the

- Art of extracting useful information from data predictions

Data Science is interdisciplinary

- Uses principles from Computer Science, Statistics, Math with domain-specific expertise of underlying application (Business, Healthcare, Bioscience, etc)

Pillars of Data Science

Data modeling

- Involves preparing /readying data for further analytics
- Processing / curation of raw data, creating structured data, extracting relevant features, and deploying appropriate inferencing

Data Analytics or Data Engineering

- Involves modeling and managing data at scale
- Data manipulation to gain insights into data
- Machine Learning plays a crucial role here

Data Visualization

- Refers to rendering the outcome of analytics
- Focused on effective communication of results
- This is a very important aspect of data science, and there are several tools out there to help (Matplotlib, Pyplot, ggPlot2 in R)

Data Ethics

- In data collection, models, and algorithms
- Pay cognizance to bias (intended and unintended)

The Cognitive Ecosystem

Cognitive / AI
Human Intelligence Exhibited by Machines

Machine Learning
"Trained" using large amounts of data & ability to learn how to perform the task

Support Vector Machine   Logistic Regression
Random Forest   Bayesian
Neural Network

Deep Learning
Many-layer neural networks

Convolution NN   Recurrent NN
Generative Adversarial Networks
Neural Turing Machines
Reinforcement Learning

Machine Learning Basics

Traditional Programming

Data → **Computer** → Output

Program

Machine Learning

Data → **Computer** → Program

Output

Machine Learning Basics

Training Data → **Computer** → Program

Training Output →

```
from __future__ import print_function
import options
import os
import sys
import utils

# Parse the command line
options.parse_args()

# only now can we import config, which depends upon args
# (also allparts, which depends upon config)
import config as cfg
import allparts
cfg.checkRequired()

# setup the environment
parwspice="pwrspice"
```

New Data → **Computer** → Output

Program →

```
from __future__ import print_function
import options
import os
import sys
import utils

# Parse the command line
options.parse_args()

# only now can we import config, which depends upon args
# (also allparts, which depends upon config)
import config as cfg
import allparts
cfg.checkRequired()

# setup the environment
parwspice="pwrspice"
```

Supervised Learning - Example

Unsupervised Learning - Example

Transaction History → *Training* → Program → **Clusters**

Recent Transaction Clusters → *Prediction* → Alert Customer?

Data Science 101 – Recommended Topics

Data Science Landscape

- Explosive growth of data
- Examples of structured and unstructured data
- Challenges & Opportunities

Data Analytics

- Machine Learning fundamentals
- Supervised vs Unsupervised learning

Data Visualization

Data Ethics and Bias reduction

What is Cybersecurity?

Cybersecurity is protecting yourself from someone stealing your digital information/personal data or from someone pretending to act as you online

Key Terms

- o **Password** – A combination of letters and numbers that is kept secret and used to gain access to a computer, website, etc.

- o **Phishing** – A scam email/website that tricks you into revealing personal information such as username, password, location, etc.

- o **Spam** – Unwanted 'junk' mail that can be used to trick you into revealing information or clicking a harmful link

- o **Virus** –Harmful "software" that attaches to other programs to hurt or destroy a computer's ability to function normally

- o **Hacker** – An unauthorized user trying to disrupt or damage a computer or network of computers

*Content was created by the IBM GBS North America Transformation Office

Imagine this ...

Someone logs into your Instagram without **YOU** knowing

They post and send messages posing as you

Your friends and followers get angry and ask about your posts or unfollow – but you have no idea what happened!

Sound annoying or scary? You can prevent it!

*Content was created by the IBM GBS North America Transformation Office

Cybersecurity - Trends

https://financesonline.com/cybersecurity-trends/

Cost of a data breach

$4.24m ↑10%
Average total cost of a data breach in 2021 study
Year-over-year growth in average total cost
Cost of a Data Breach Report 2021
IBM Security

20% $4.37m
Share of breaches caused by compromised credentials
Average cost of a compromised credentials breach
Cost of a Data Breach Report 2021
IBM Security
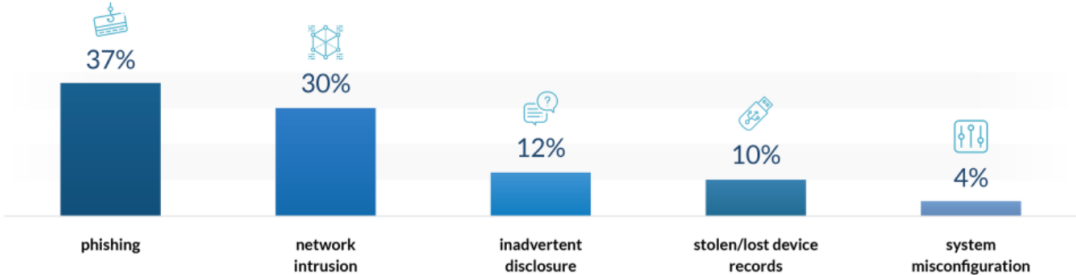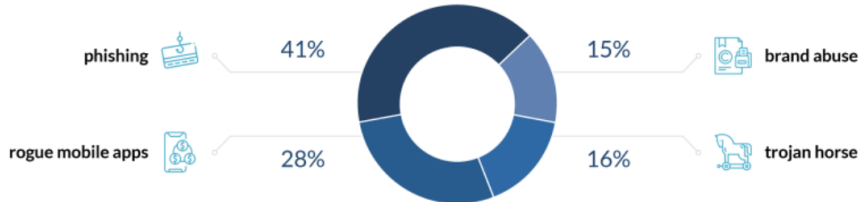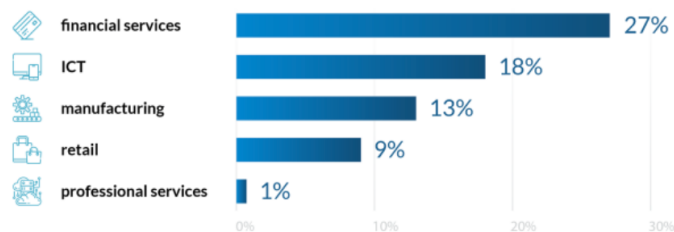
https://www.ibm.com/security/data-breach

Cybersecurity 101 – Recommended Topics

Cybersecurity landscape

- What is it and Why is it important
- Cost of a data breach

Cybersecurity basics

- Fundamentals, Types of Attacks
- CIA (Confidentiality, Integrity, Availability) Triad

Protection

- Identity, Assets, Data, Apps
- Encryption, Cryptography

Laws, Governance, Compliance

What is Cloud Computing?

Cloud computing transforms IT infrastructure into a utility

by delivering computing services over the internet

These services include Software, Databases, Servers, and Storage

Users normally pay rent for these services

Compared to "build your own infrastructure model" cloud computing provides
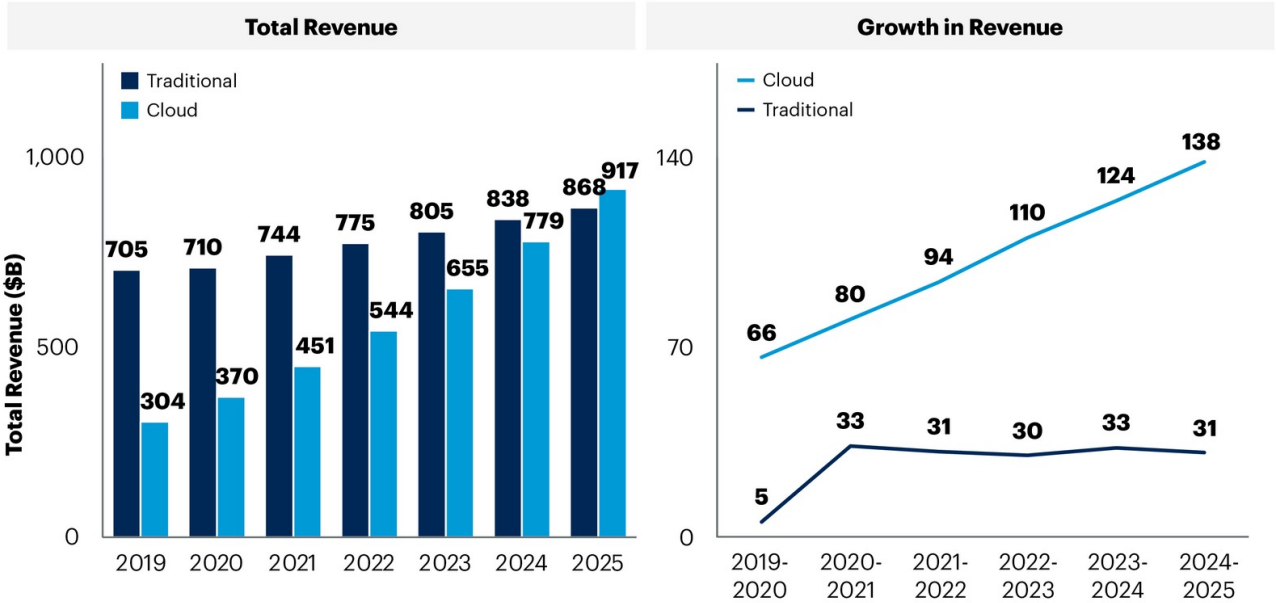
Easier set-up and availability of software / support

Lower costs compared to dedicated infrastructure

Flexibility to scale-up or scale-down based on need

Cloud Growth Forecast

Figure 1: Sizing Cloud Shift, Worldwide, 2019 – 2025

Cloud Service Models
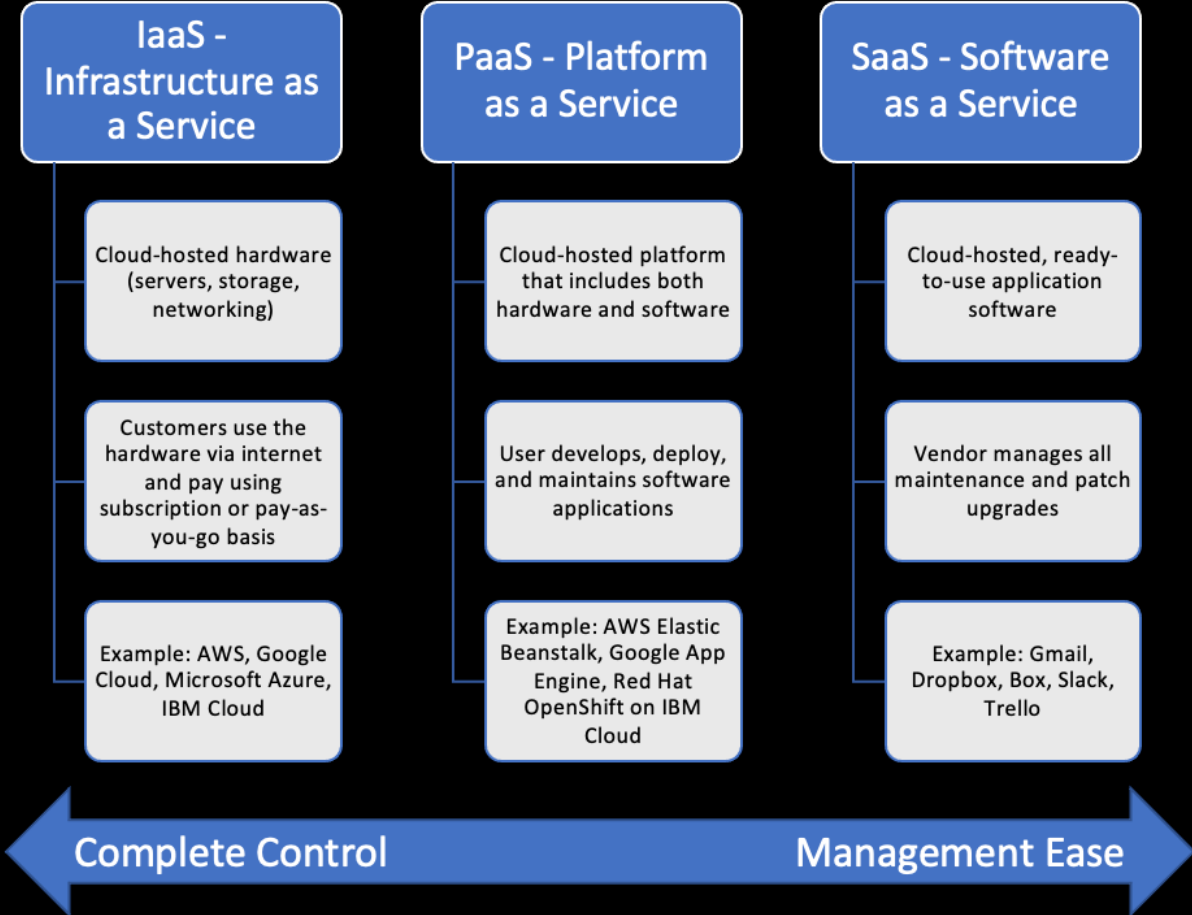
Cloud Service Models – a closer look

IaaS - Infrastructure as a Service

- Cloud-hosted hardware (servers, storage, networking)
- Customers use the hardware via internet and pay using subscription or pay-as-you-go basis
- Example: AWS, Google Cloud, Microsoft Azure, IBM Cloud

PaaS - Platform as a Service

- Cloud-hosted platform that includes both hardware and software
- User develops, deploy, and maintains software applications
- Example: AWS Elastic Beanstalk, Google App Engine, Red Hat OpenShift on IBM Cloud

SaaS - Software as a Service

- Cloud-hosted, ready-to-use application software
- Vendor manages all maintenance and patch upgrades
- Example: Gmail, Dropbox, Box, Slack, Trello

◄ Complete Control   Management Ease ►

Cloud Computing Services: Who Manages What?

	Traditional IT	IaaS	PaaS	SaaS
Applications	You manage	You manage	You manage	Provider manages
Data	You manage	You manage	You manage	Provider manages
Runtime	You manage	You manage	Provider manages	Provider manages
Middleware	You manage	You manage	Provider manages	Provider manages
OS	You manage	You manage	Provider manages	Provider manages
Virtualization	You manage	Provider manages	Provider manages	Provider manages
Servers	You manage	Provider manages	Provider manages	Provider manages
Storage	You manage	Provider manages	Provider manages	Provider manages
Networking	You manage	Provider manages	Provider manages	Provider manages

You manage / Provider manages

https://www.ibm.com/cloud/learn/iaas-paas-saas

Cloud Deployment Models

Public Cloud

- Third-party service provider makes computing resources
- Available to everyone with an internet
- Lower cost (sometimes free), convenient, easily scalable

Rent an apartment

Private Cloud

- Cloud infrastructure operated exclusively for a set of users (or a company)
- Highly flexible and provides greater control over resources
- Higher Cost (especially upfront)

Own a single-family home

Hybrid Cloud

- Integrates both Public and Private Cloud Models
- Tighter control over sensitive data while lower costs for less critical data
- Allows for easy movement of workloads between both cloud models

Own a condo, but pay for maintenance

Cloud 101 – Recommended Topics

Cloud Computing Landscape

- History, Introduction to Cloud Computing

Fundamentals of Cloud Computing

- Cloud service modes and types
- Virtualization, Hypervisors
- Security and challenges

Cloud Hands-on

- Building and deploying your application on the cloud

IBM Resources available to academia

FREE

- https://www.ibm.com/academic/home

- https://www.ibm.com/academic/topic/data-science

- https://www.ibm.com/academic/topic/security

- https://www.ibm.com/academic/topic/cloud

IBM | **IBM Academic Initiative** | Topics ⌄ | Usage terms ⌄ | Additional Resources | FAQ

IBM Academic Initiative

Harness the power of IBM. Get easy no-charge access to the tools you need to develop the next great thing. Enjoy powerful technical and strategic resources from IBM. Jump right in with access to powerful services and the most prominent open-source computer technologies, or take advantage of hands-on resources that will teach you about data science, artificial intelligence, security and more.

Already registered? Log in | Register now

Q & A