



Center of Excellence: How Micro-Electronics Revolutionizing the World

by: Tassadaq Hussain

Professor Namal University Mianwali

Director Centre for AI and BigData

PakASIC.com

Collaborations:

Barcelona Supercomputing Center, Spain

European Network on High Performance and Embedded Architecture and Compilation

Pakistan Supercomputing Center

Introduction



Education:

PhD. Barcelona-Tech
Microsoft Research, Infineon
Technologies France, Microsoft
Research Cambridge, IBM

Suspenseful record of academic
management as Professor and Dean

Enhanced Education Quality by
Inculcating Outcome Based
Education by Applied and
Sustainable Projects

Experience:

19+ year's versatile experience in the area
of Computer Architecture, AI, Software
Architecture, Big-Data Architecture
Served National and International Academia,
Industry and Government

- Barcelona Science Park Spain
- Cambridge Science Park UK
- Technopolis Of Sofia-Antipolis, France



Innovation, Research and Commercialization



Innovation and Research

- 110+ Million Pkr National and Int'l Funding.

Supercomputing and Artificial Intelligence

Smart Electric Motor Controllers

Biomedical Applications

- 80+ Publications
- 10 Patents
- 10 MVPs
- 5 Int'l Collaborations

Development & Commercialization

60+ Million of Industrial Investments.

Developed Digital Systems for Industry.

Transform Idea into product.

Innovation and Commercialization for Sustainable economic and industrial development.

Capacity Building:

Conducted more than 50 national and international workshops and training on Commercializable research, Writing successful grant proposal, and research and innovation.

Provides Consultancy and Support for Entrepreneurship, Start-ups, Business Innovation and Technology transfer.



PAKISTANTM
SUPERCOMPUTING



EntrepreneurialPark
Research Innovation Commercialization



Foot Analytic



ComputingPark

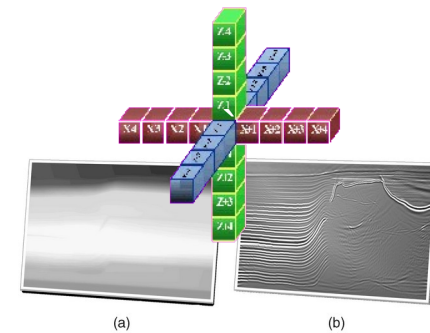
Int'l Projects

- Design Ultra Low Cost Display Camera Interface for Mobile Baseband XGold Chip (**Infineon Technologies, 200 million single chip**)



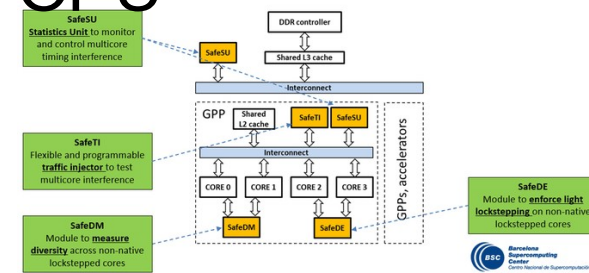
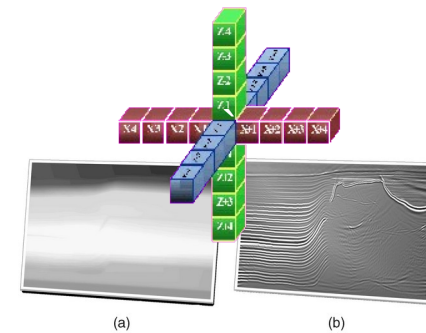
Int'l Projects

- Design Ultra Low Cost Display Camera Interface for Mobile Baseband XGold Chip (**Infineon Technologies, 200 million single chip**)
- Implementation of Reverse Time Migration on FPGAs (**BSC-REPSOL, PLDA Italia, Cambridge Science Park**)



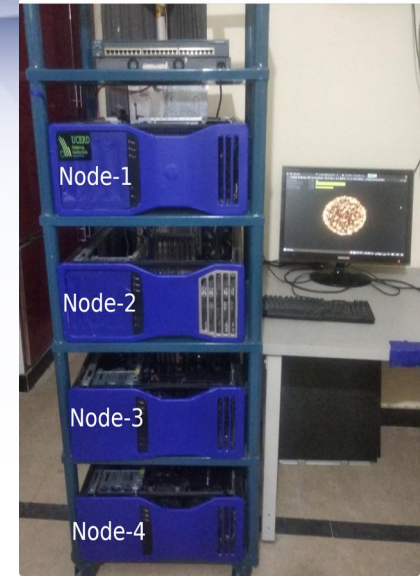
Int'l Projects

- Design Ultra Low Cost Display Camera Interface for Mobile Baseband XGold Chip (**Infineon Technologies, 200 million single chip**)
- Implementation of Reverse Time Migration on FPGAs (**BSC-REPSOL, PLDA Italia, Cambridge Science Park**)
- Open source European full-stack ecosystem based on a new RISC-V CPU (**Barcelona Supercomputing Center**)



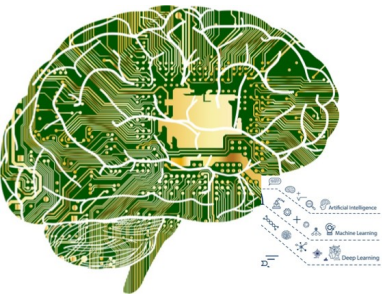
National

- Supercomputing and AI for Health Sciences
- Pakistan Supercomputing Center
- FPGA Power Supercomputer
- Scalable Heterogeneous Supercomputing System
- Smart Motor Controller
- FootAnalytic
- VR/AR for Rehabilitation
- Live-Stock Breed Identification System



Objectives of the talk

- World Data Size = 130 Zettabytes, doubling every 18 months.
- To handle big-data, AI algorithms are the only solution.
- The computational demands of AI algorithms are experiencing exponential growth. (ExaFLOPS/Day)
- Micro-Electronics is the only solution to store big-data and process the AI.



Secure
Reliable
Programmable
Customize-able
Indigenous

(Till 2021, 200 Billion CPU cores in the world running)

Data

- Raw facts

Information

- Processed data with contextual meaning
- Tells us "who" "what" "where" "when"

Knowledge

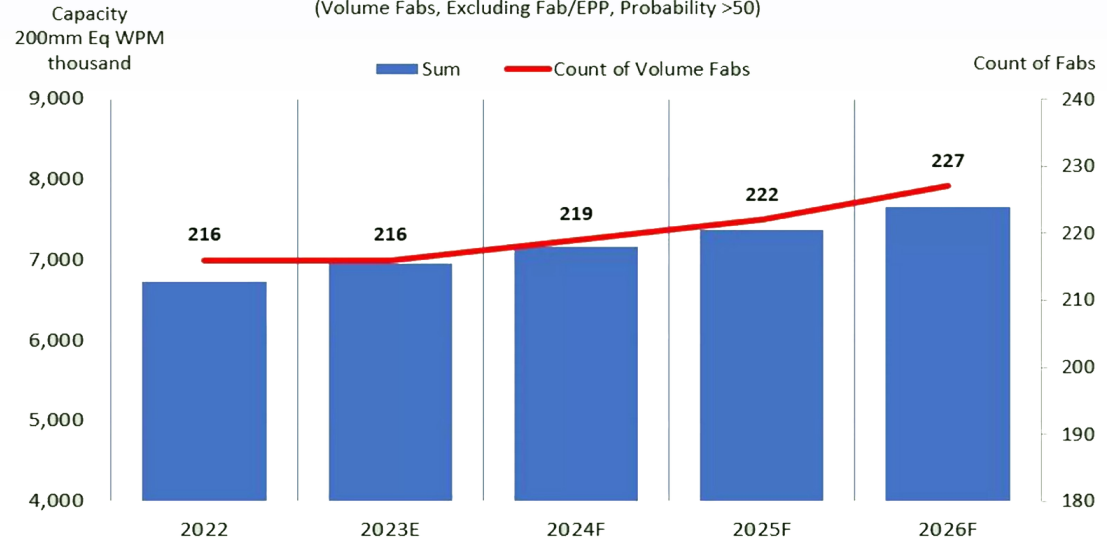
- Collective information gained through individual expertise
- Describes "how"

Wisdom

- Knowledge that has an evaluative component
- Describes "why"

200mm Semiconductor Capacity and Fab Counts

(Volume Fabs, Excluding Fab/EPP, Probability >50)



200mm Fab Outlook to 2026, 3Q23 Update, Published by SEMI

7.7 million wafers per month (WPM)

Mastery of Chip Design is essential; a lack may result in unforeseen consequences.

Objectives

Educate , Collaborate and Accelerate

The goal of this school is to foster **interdisciplinary collaboration** and **teamwork across departments** within the **University** through the exploration of Micro-electronics, and Intelligence applicatoins.

by:

Leveraging the collective expertise and resources, challenges and opportunities

for:

Advancing research, education, and societal impact.

“The future we will “invent” is a choice we make jointly, not something that happens.” Jordi

Previous Summer School on Supercomputing for AI and BigData and Chip Design

https://github.com/ucerd/Summer-School-2023_1

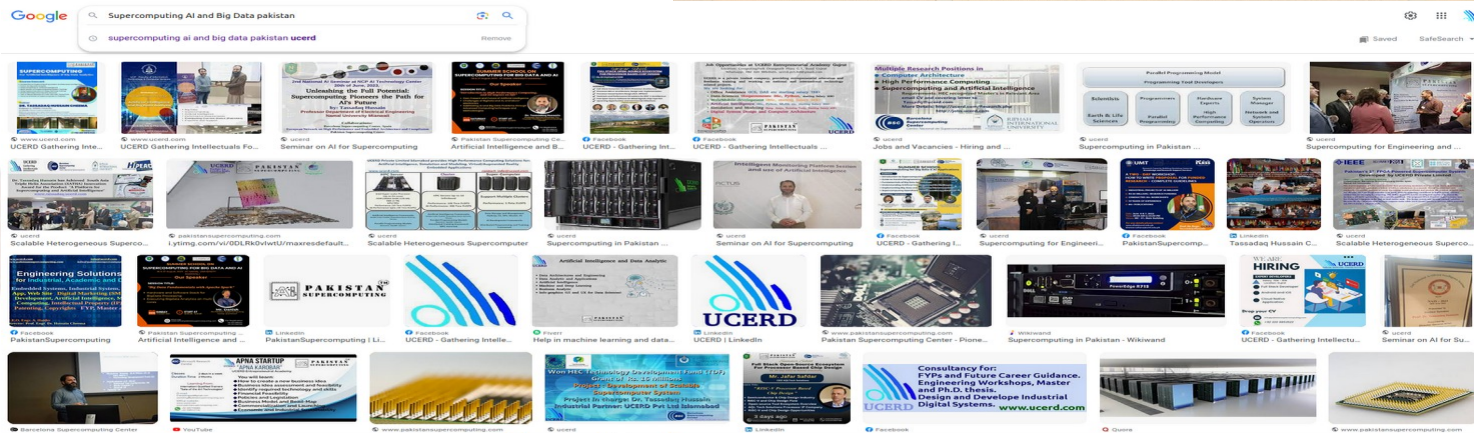
16 Speakers (20 Sessions)


Github: More than 1000 download in 3 weeks

3 National Collaborations

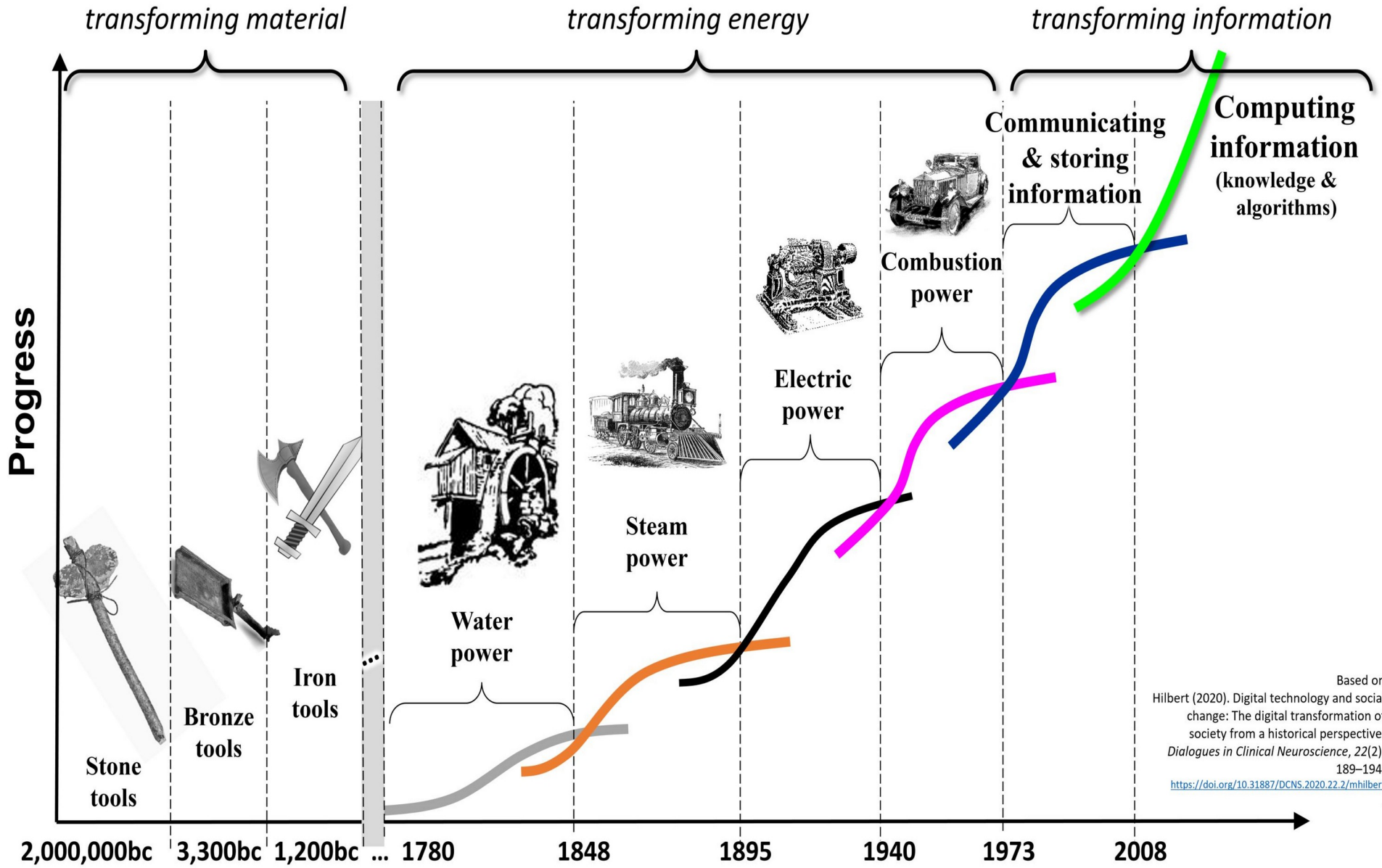
2 Linkages (Projects)

Towards International Recognition (PRACE)



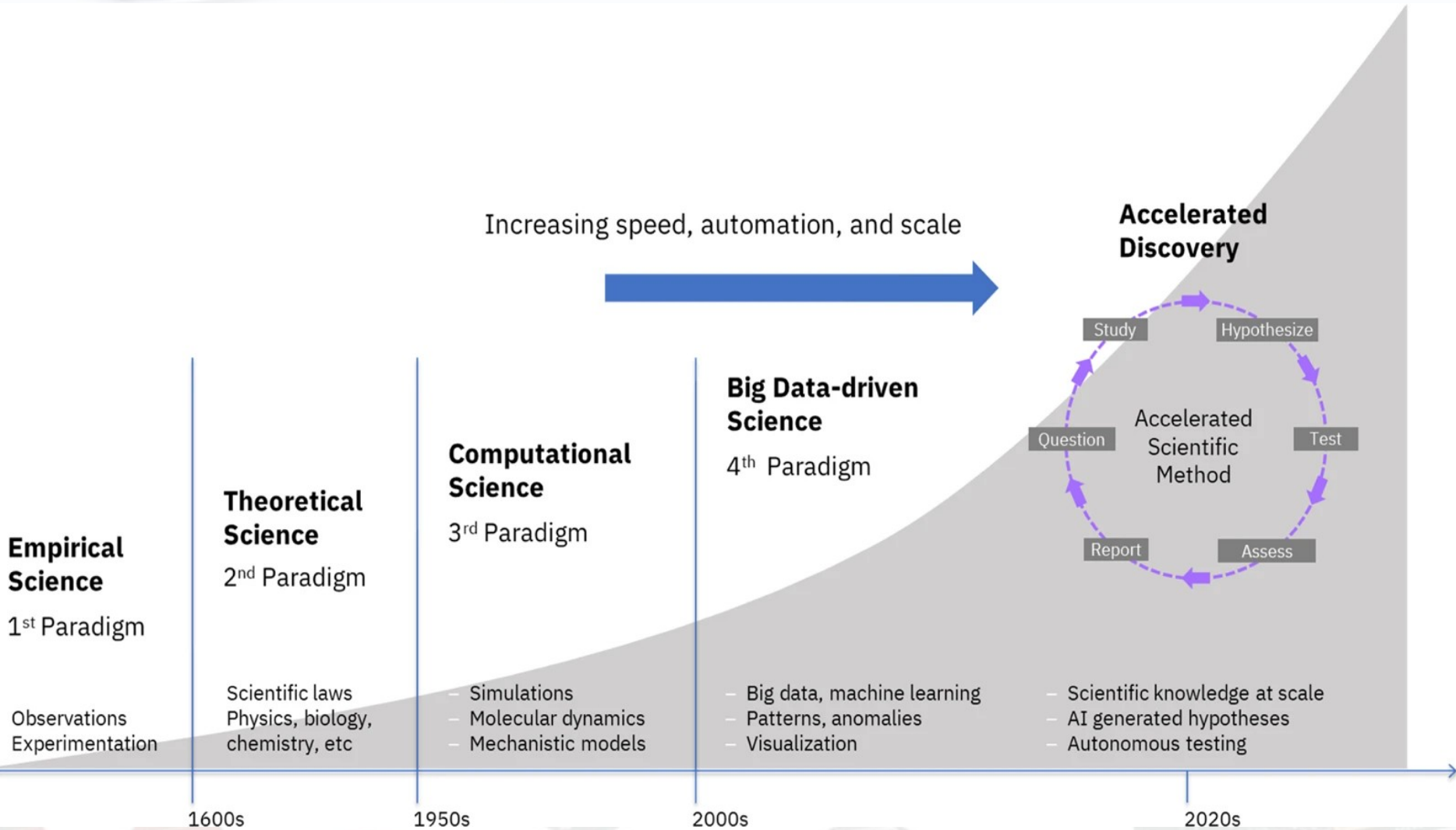
- 
- **Mankind Progress and Industrial Revolution**
 - Age of Big Data and AI
 - Micro-electronics! Revolutionizing the World
 - Namal Chip Design and HPC Facility

Mankind Progress



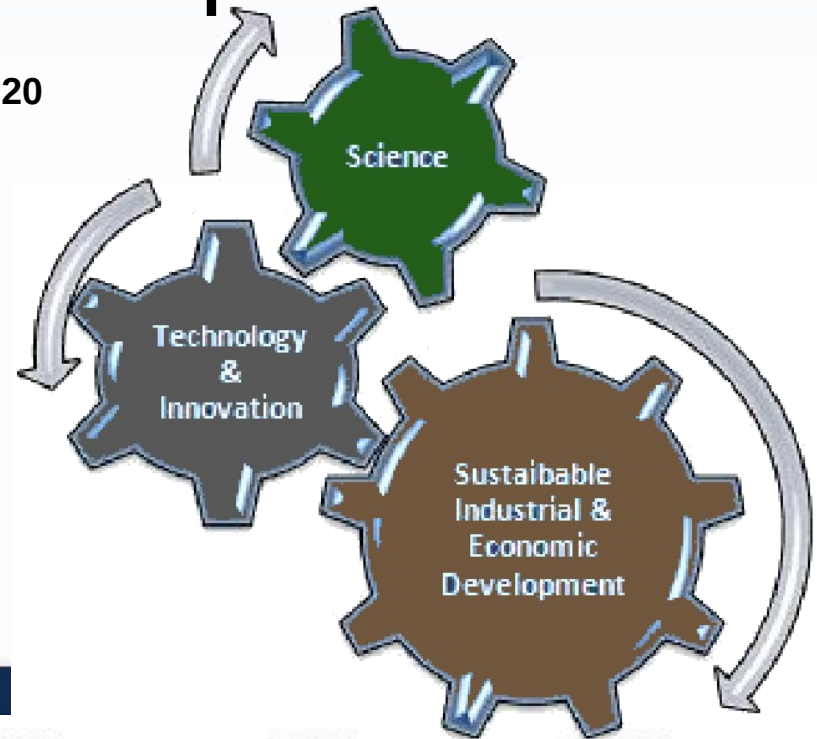
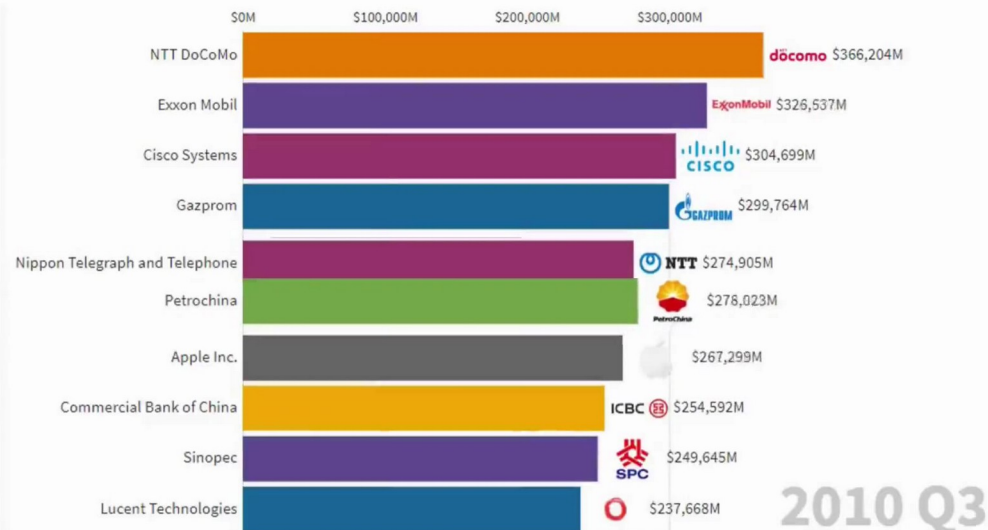
Based on
Hilbert (2020). Digital technology and social
change: The digital transformation of
society from a historical perspective.
Dialogues in Clinical Neuroscience, 22(2),
189–194.
<https://doi.org/10.31887/DCNS.2020.22.2/mhilbert>

From Age of Empirical Science to Data-Science

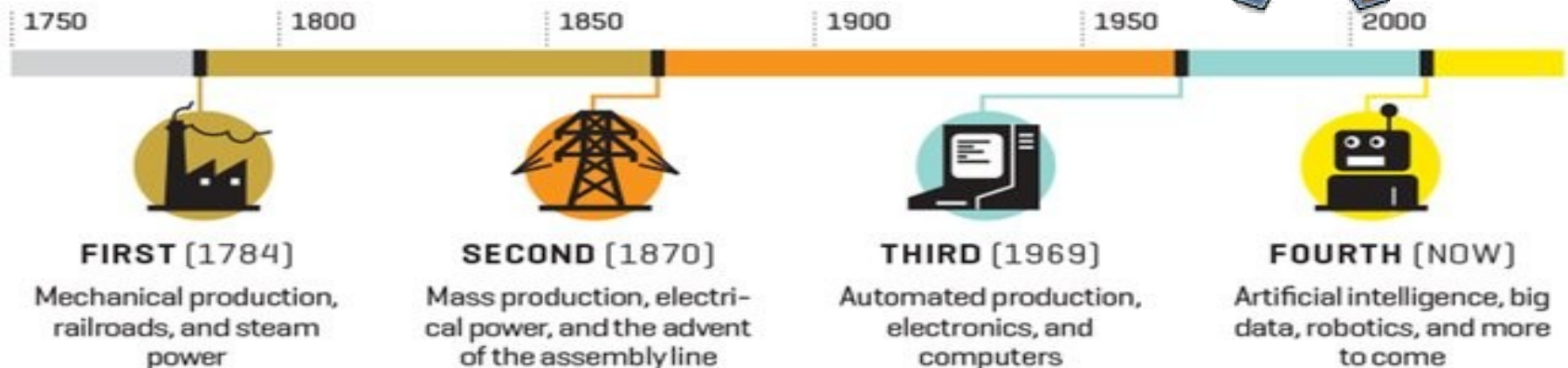


Industrial Revolutions and Sustainable Developments

Top 10 Biggest Companies By Market Size From 2010 - 2020



PakistanSupercomputing.com



Ecosystem of Modern Industry



Life Science



Earth Science



Social Science

Science

175 ZByte @2025

80%
Data-Sciences

Data

100 ExaFLOPS
@2020

87.04 B\$
234.6 B\$ @2025

AI

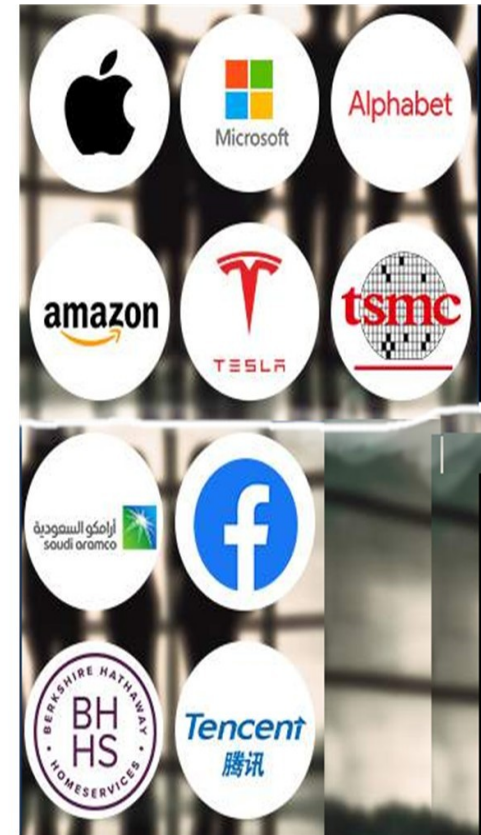
Top500 List
8 PetaFLOPS
@2022

uProcessor
100 B\$ @2020

30% Cell Phone
20% Embedded
App
50 Servers, PCs etc.

Computing

Digital Industrial Age
5.5 Trillion \$ Revenue@2021

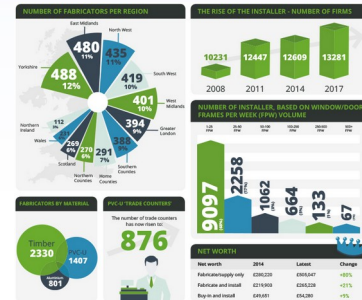


Four Tiers of Digital Industry

Tier 1:

Front End Development Industry (Web, Infographics etc)

Tens to Hundred of billions dollars industry



Tier 2:

Data Management Industry (Analytics Classification, etc)

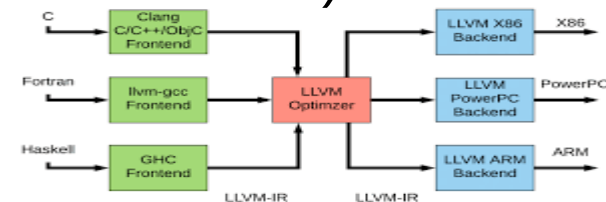
Under Hundred of billion dollars industry



Tier 3:

Software Development (Compiler, AI Models, Applications etc.)

Over Hundreds of Billions of dollars industry




Tier 4:

Hardware Development (Semiconductor, etc)

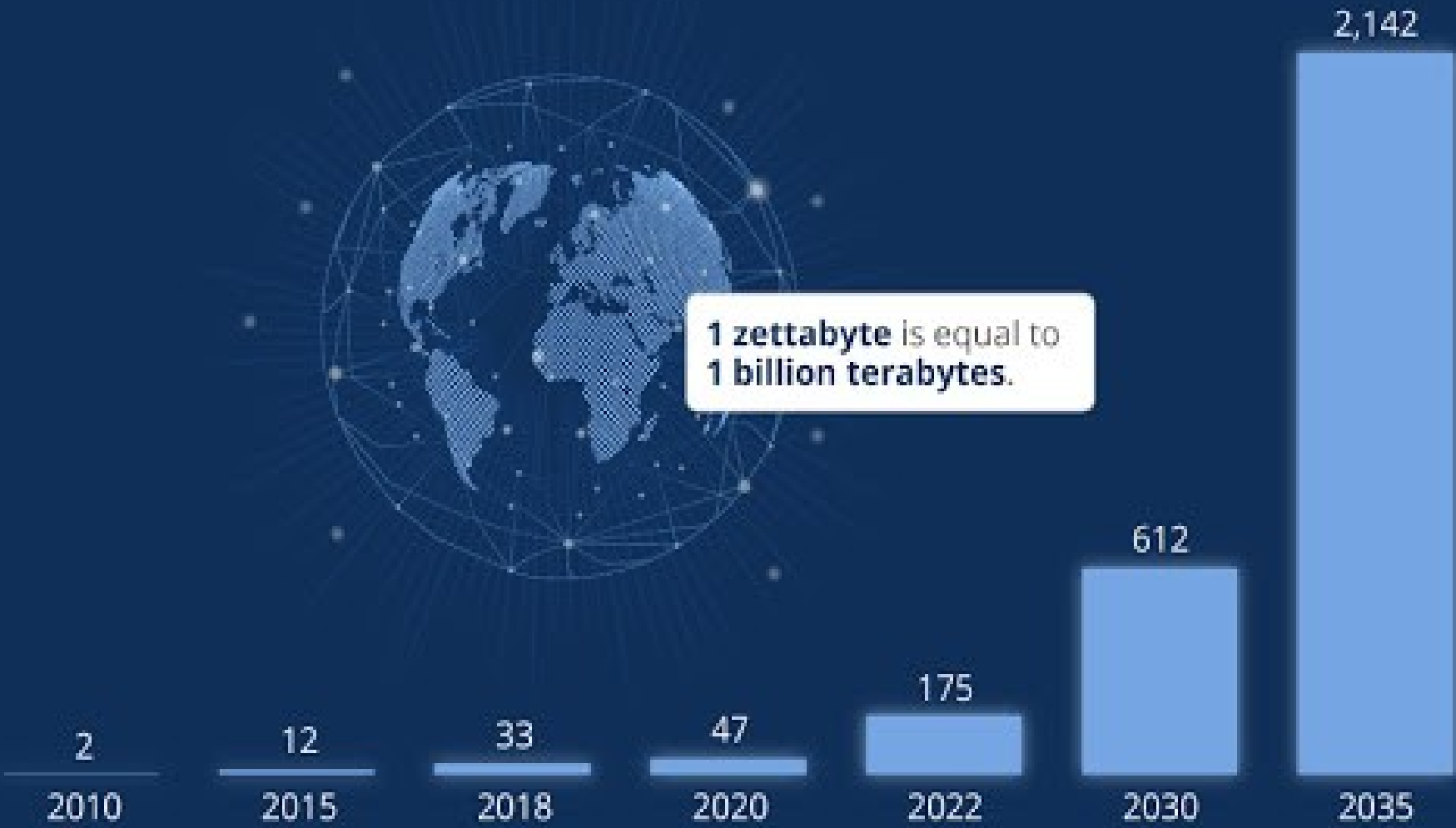
Hundreds of billions to over a trillion dollars Industry



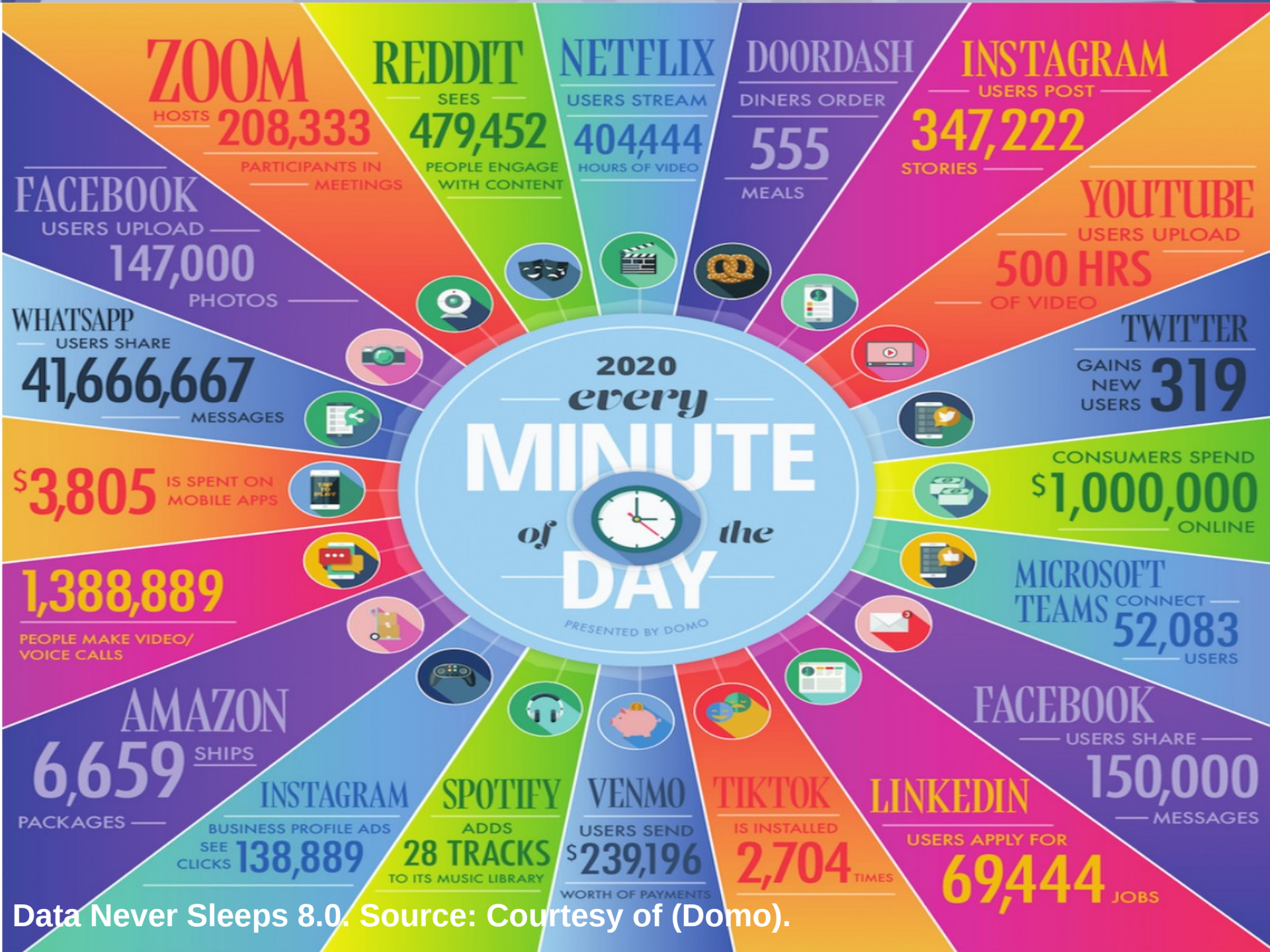
- 
- Mankind Progress and Industrial Revolution
 - **Age of Big Data and AI**
 - Micro-electronics! Revolutionizing the World
 - Namal Chip Design and HPC Facility

Global Data Creation is About to Explode

Actual and forecast amount of data created worldwide 2010-2035 (in zettabytes)



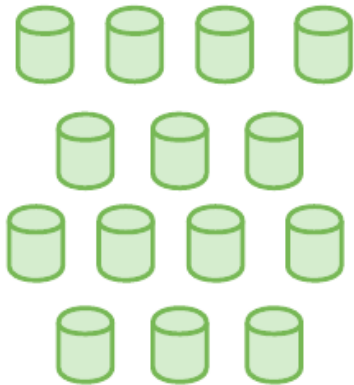
57.76 US\$



Data Never Sleeps 8.0. Source: Courtesy of (Domo).

Types of Data and its Challenges

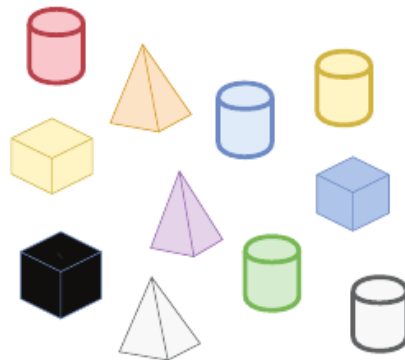
Data at rest



Terabytes to zettabytes
of data to process

Volume

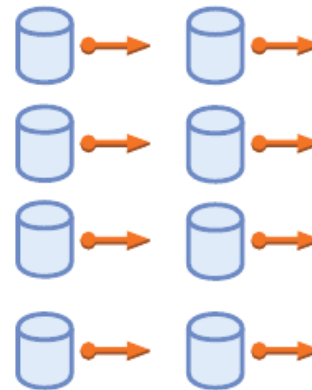
Data in many forms



Structured,
unstructured, and semi-
structured

Variety

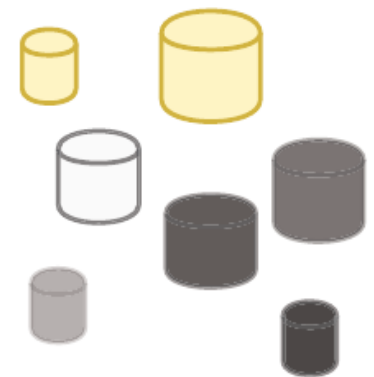
Data in motion



Streaming data,
microseconds to seconds
to respond

Velocity

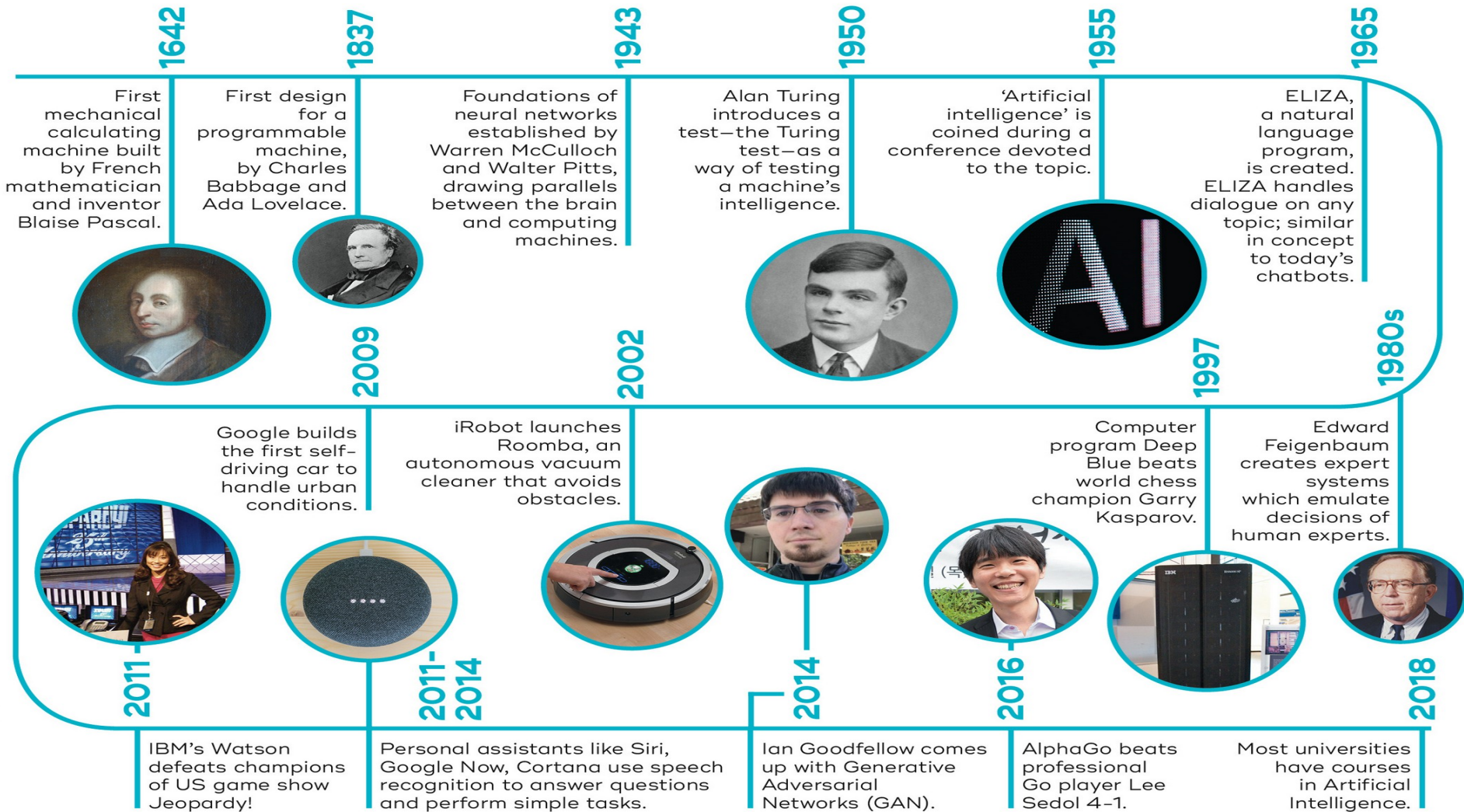
Data in doubt



Uncertainty due to data
inconsistency, ambiguities,
deception, and model
approximations

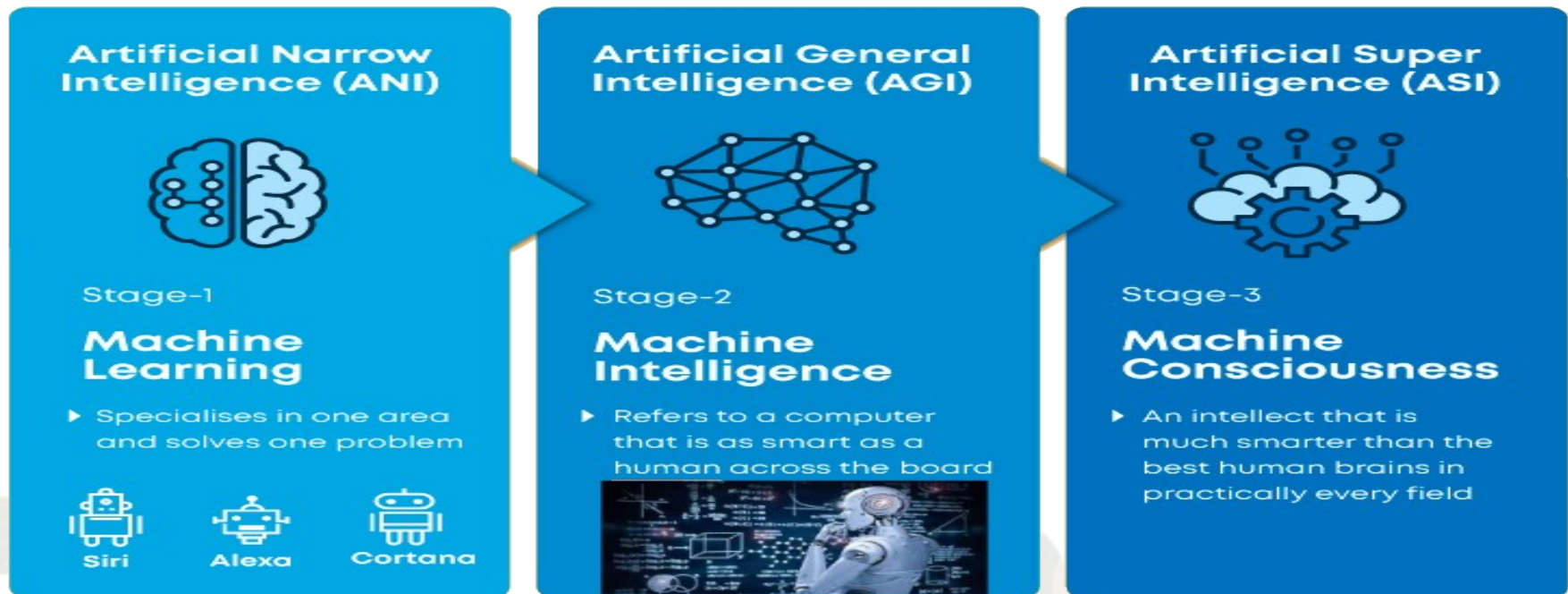
Veracity

AI: The Only Solution for BigData



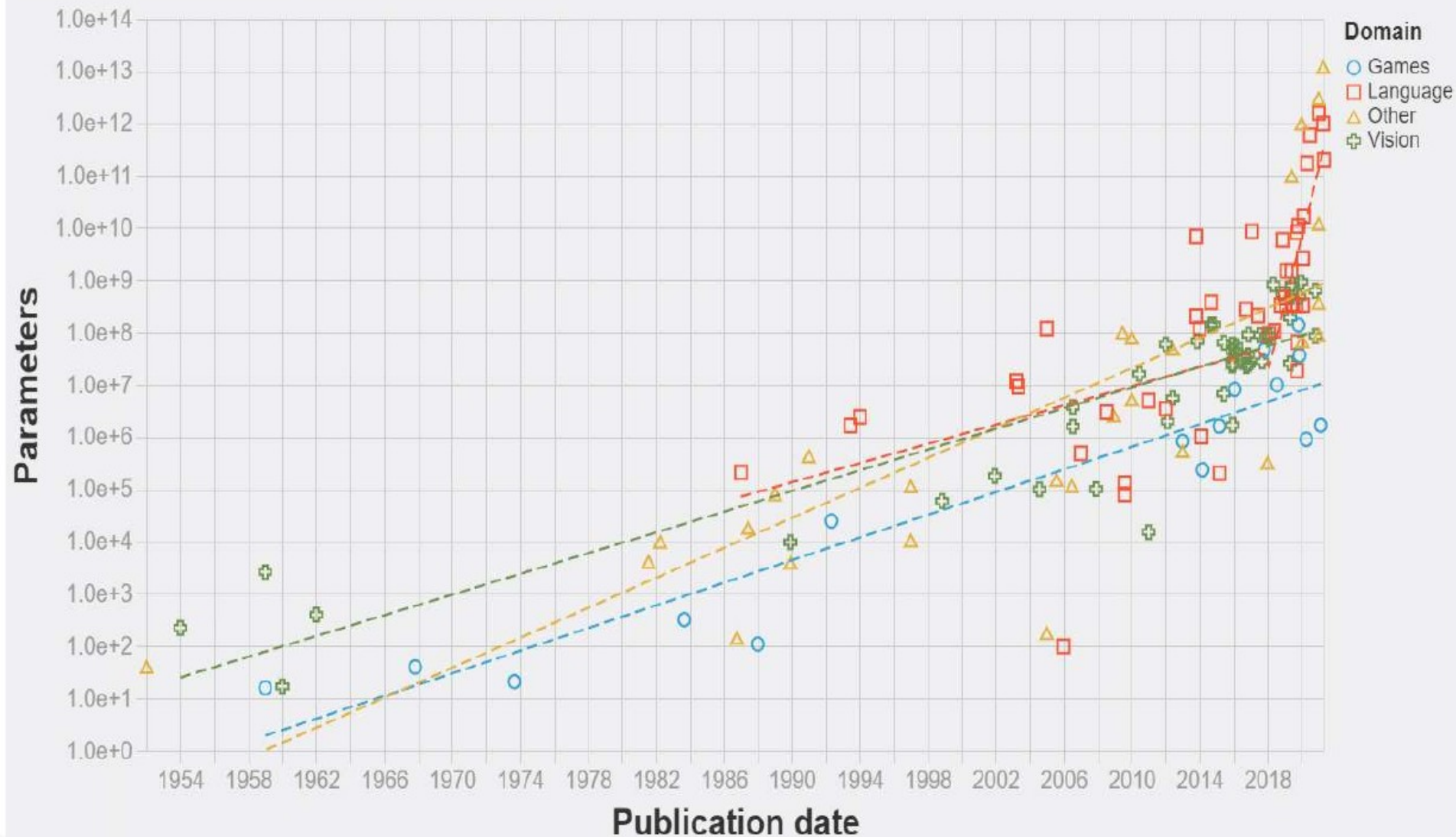
BigData and AI Algorithms

- **Performance**
 - **Execution Time**
 - **Accuracy** “The accuracy of the model is inherently tied to the quality, diversity, and representativeness of the data used for training and evaluation.”
 - **Scalability** “Methods that scale with computation are the future of Artificial Intelligence” — Rich Sutton,



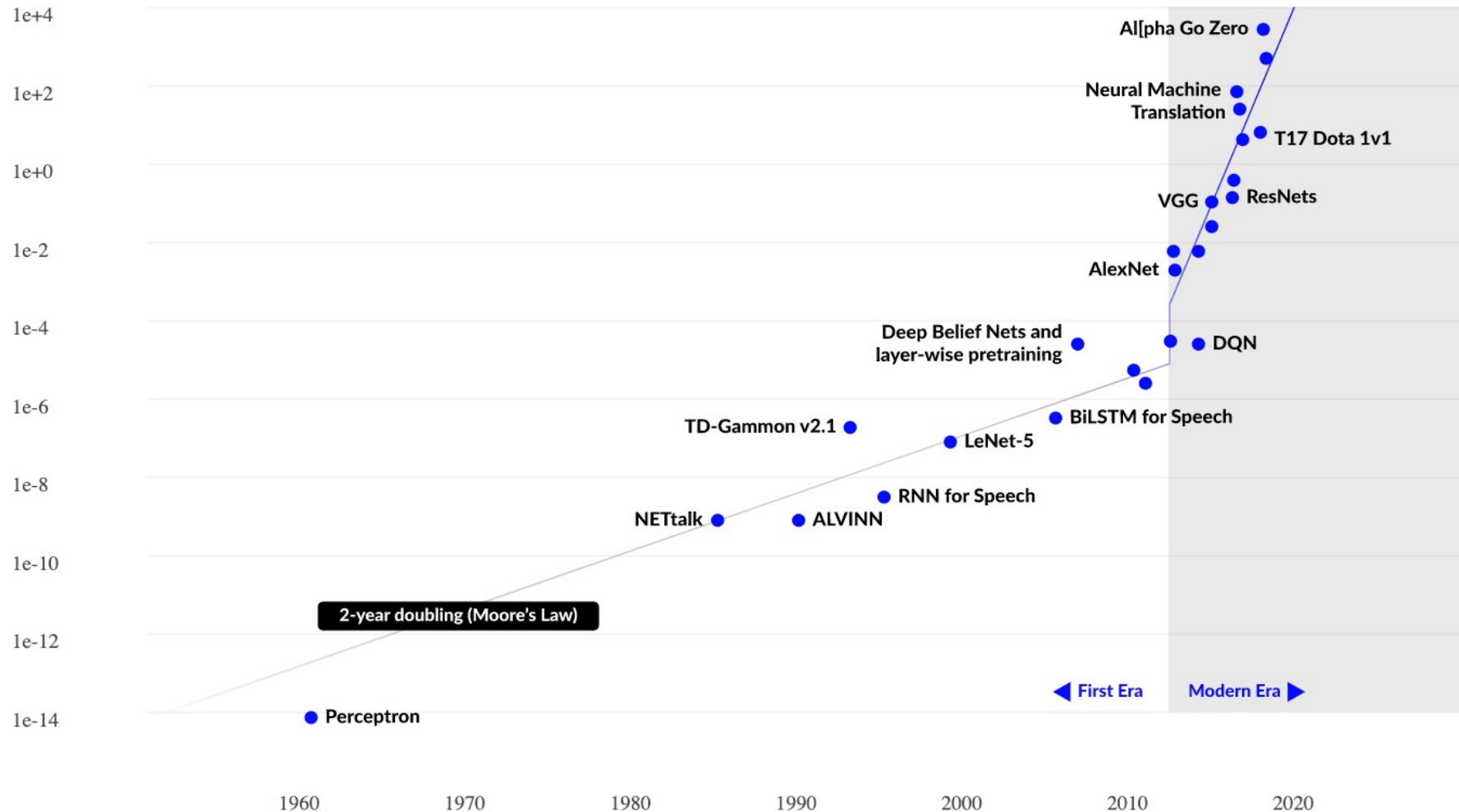
DL Relentless growth in model size

Parameter count of ML systems through time

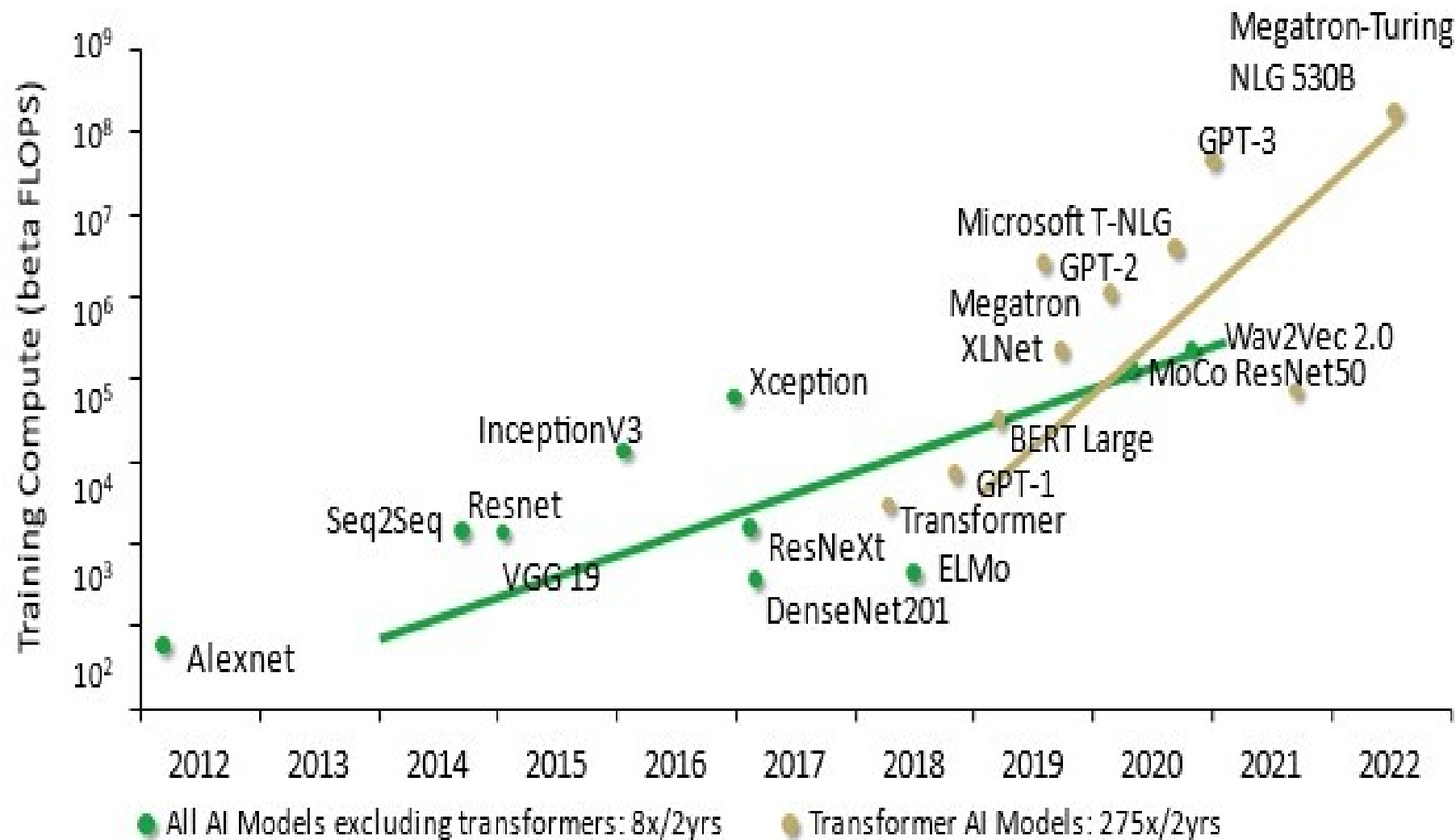


Computation Demand

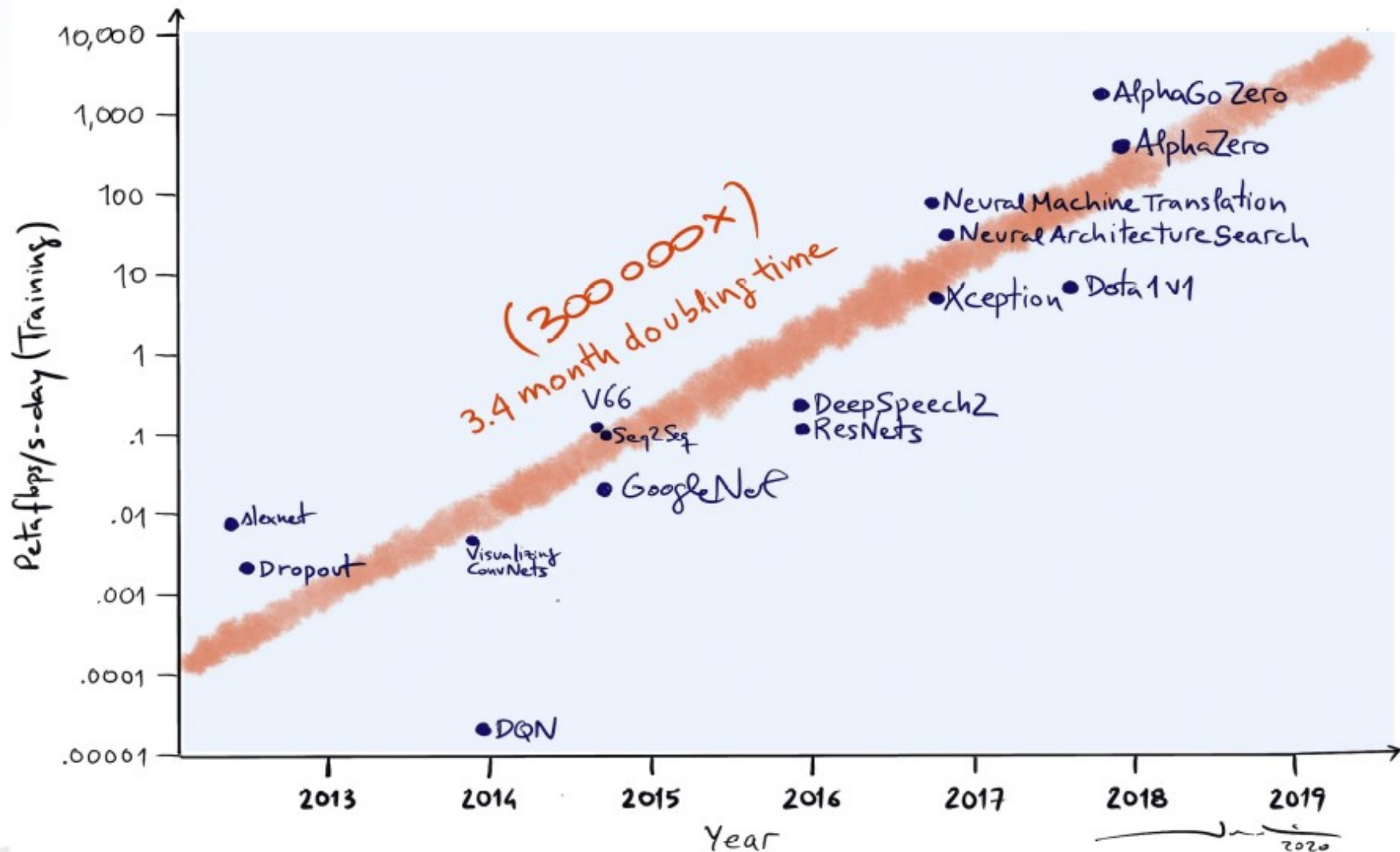
Petaflop/s-days




The total amount of compute, in petaflop/s-days,[2] used to train selected results that are relatively well known, used a lot of compute for their time, and gave enough information to estimate the compute used.



AI Computational Requirements



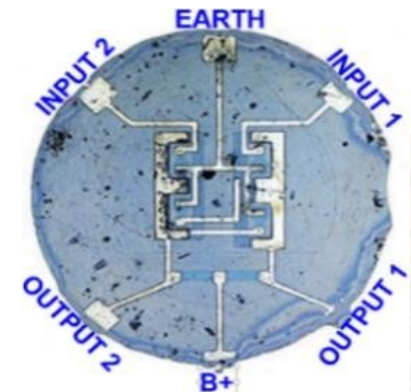
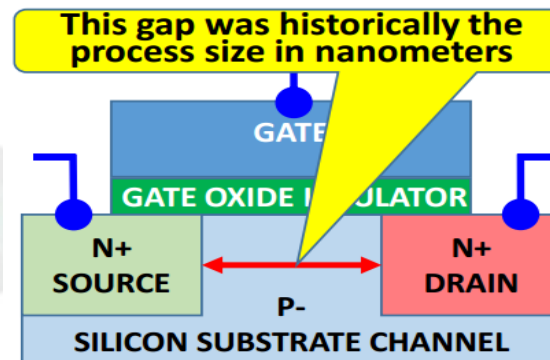
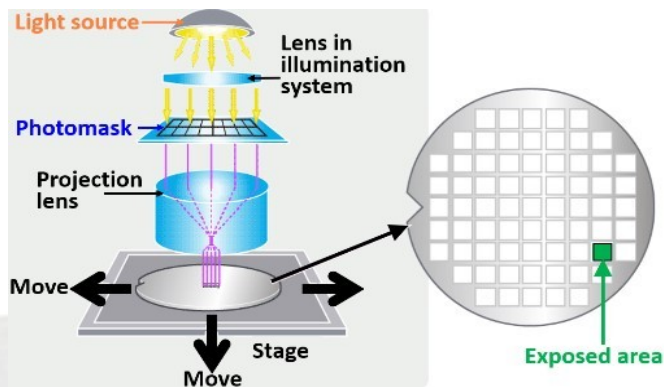
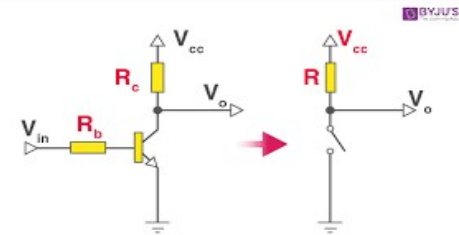
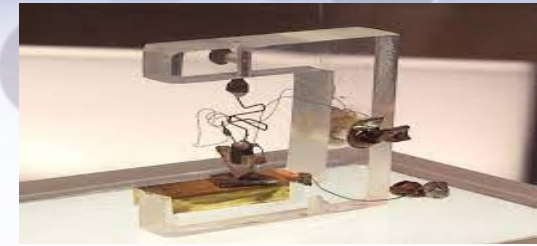
- 
- Speaker Introduction
 - Objectives of this Event
 - Mankind Progress and Industrial Revolution
 - Age of Big Data and AI
 - **Micro-electronics! Revolutionizing the World**
 - Namal Chip Design and HPC Facility

Democratization in Microelectronics

- GCC has **revolutionized** the software industry.
- Linux has **revolutionized** computing industry.
- Arduino has **revolutionized** embedded computing.
- Mathematical Models, Development Frameworks and Opensource datasets have been **revolutionizing** the computing intellectuality.
- RISC-V is **revolutionizing** the Secure Computing.
- Open Silicon is next => Indigenous Development.

History of Transistor

- **Transistor:** Key invention of the last century
- Until the late 1950s, computer circuits comprised discrete components like transistors, resistors, diodes, and capacitors soldered by hand on circuit boards.
- Transistorized computers were large, power-hungry, and had complex wiring due to individual transistor connections.
- In 1959, Fairchild Semiconductor's Robert Noyce and Shockley introduced a breakthrough with silicon integrated circuits (ICs).



Birth of Computing

Mechanical Computing by Charles Babbage

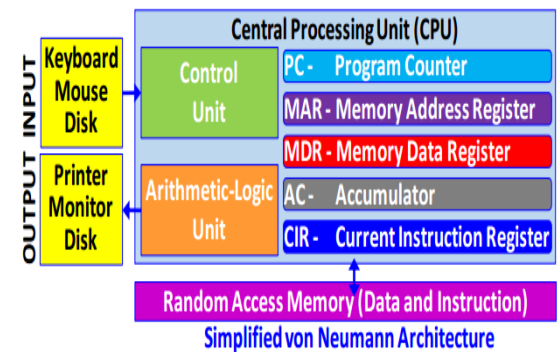
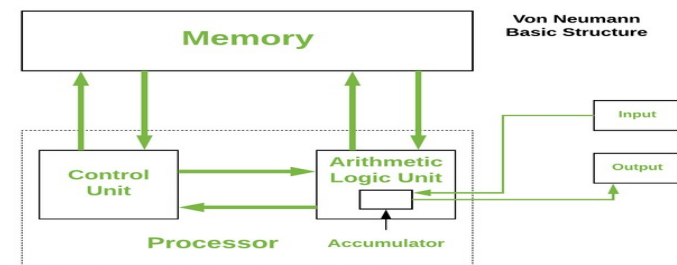
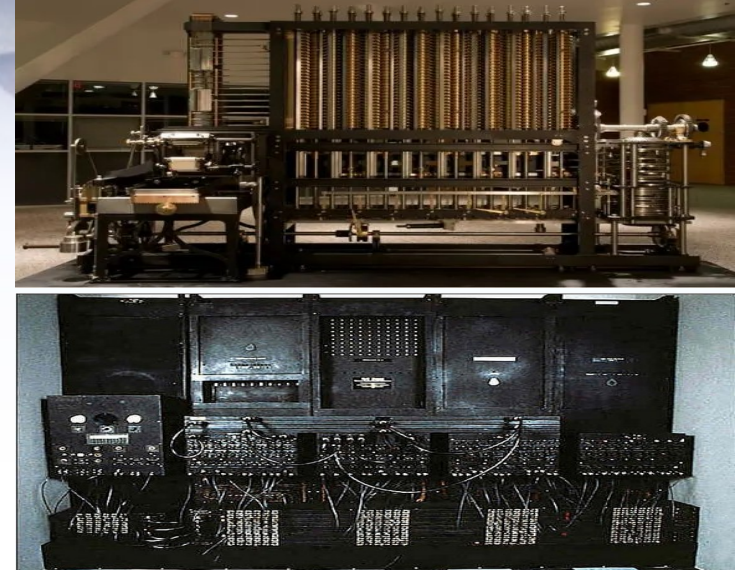
- Arch: gears, levers, and rotating shafts.
- Storage: 1K Decimal Digits
- Programming: Punch Card
- Output: Printer

Digital Computing John Von Neumann

- 1945 Electronic Discrete Variable Automatic Computer
- 1946 ENIAC (Electronic Numerical Integrator and Computer)
- **Arch:** 17,468 vacuum tubes, 7,200 crystal diodes, 1,500 relays, 70,000 resistors, 10,000 capacitors, and around 5 million hand-soldered joints.
- **Performance:** 5,000 additions or 357 multiplications per second. @100KHz Clk
- "Fixed program" computer with switches and plug boards
- Input and Output: Data was input using punched cards and output through various display devices, including card punchers, printers, and oscilloscopes.

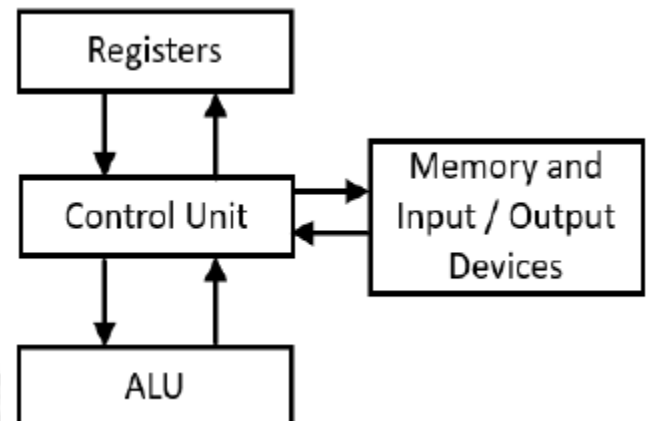
IBM PC 1981.

- 8088 Processor Architecture
- 4.77 MHz Clock, 16KB RAM

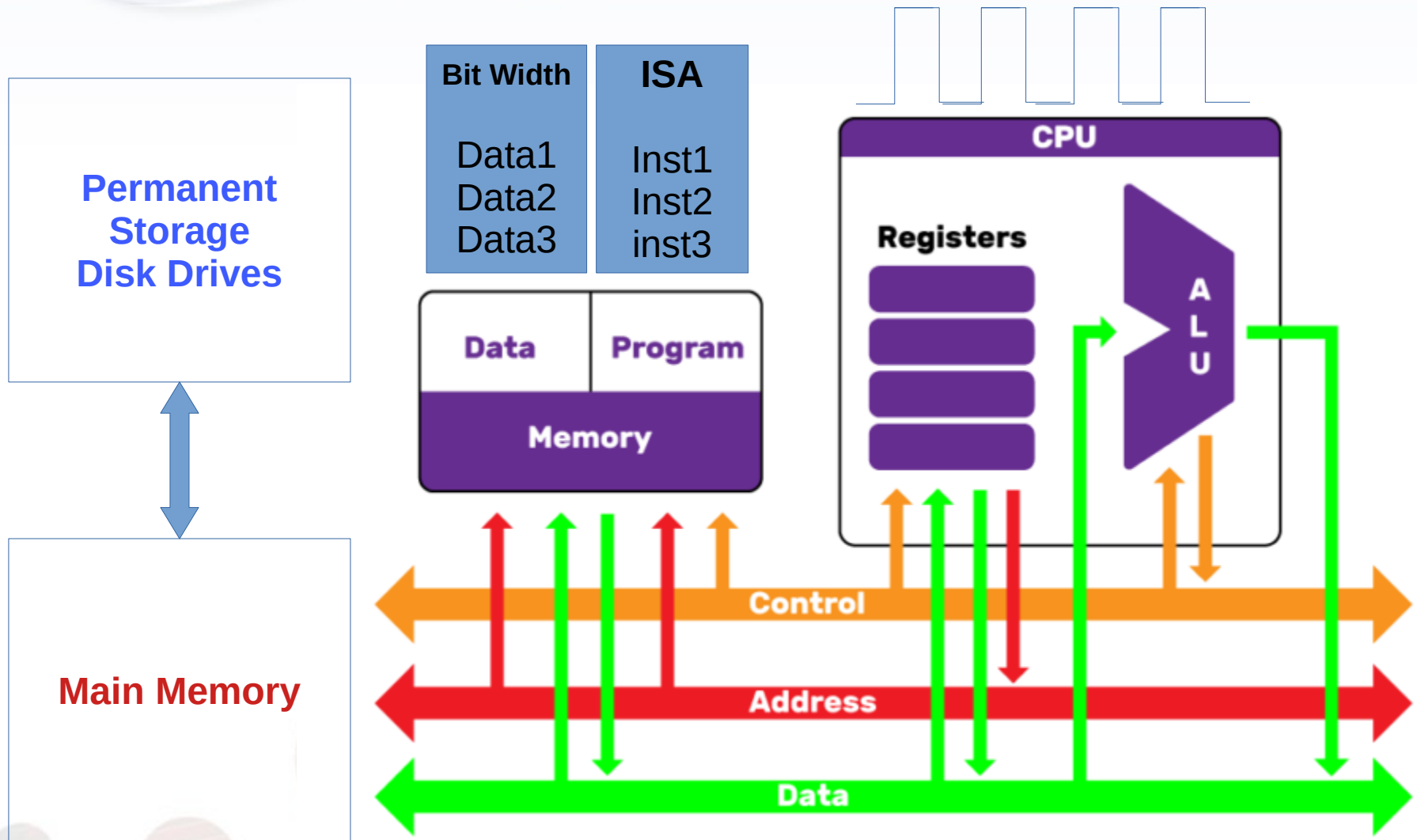


Basic Processor Architecture

- A central processing unit (CPU) gets instructions and/or data from memory, decodes the instructions and then ***sequentially*** performs them.
- Memory is used to store both program and data instructions
 - Program instructions are coded data which tell the computer to do something
 - Data is simply information to be used by the program



Information and Computer



Microprocessor Development Directions

Increasing of clock frequency and speed instruction stream processing

Processing of large collection of data in single processor instruction - SIMD

Control path multiplication – multi threading

RISC processors

- MIPS
- IBM Power4
- Alpha
- RISCV

CISC processors

- IA32
- AMD x86-64

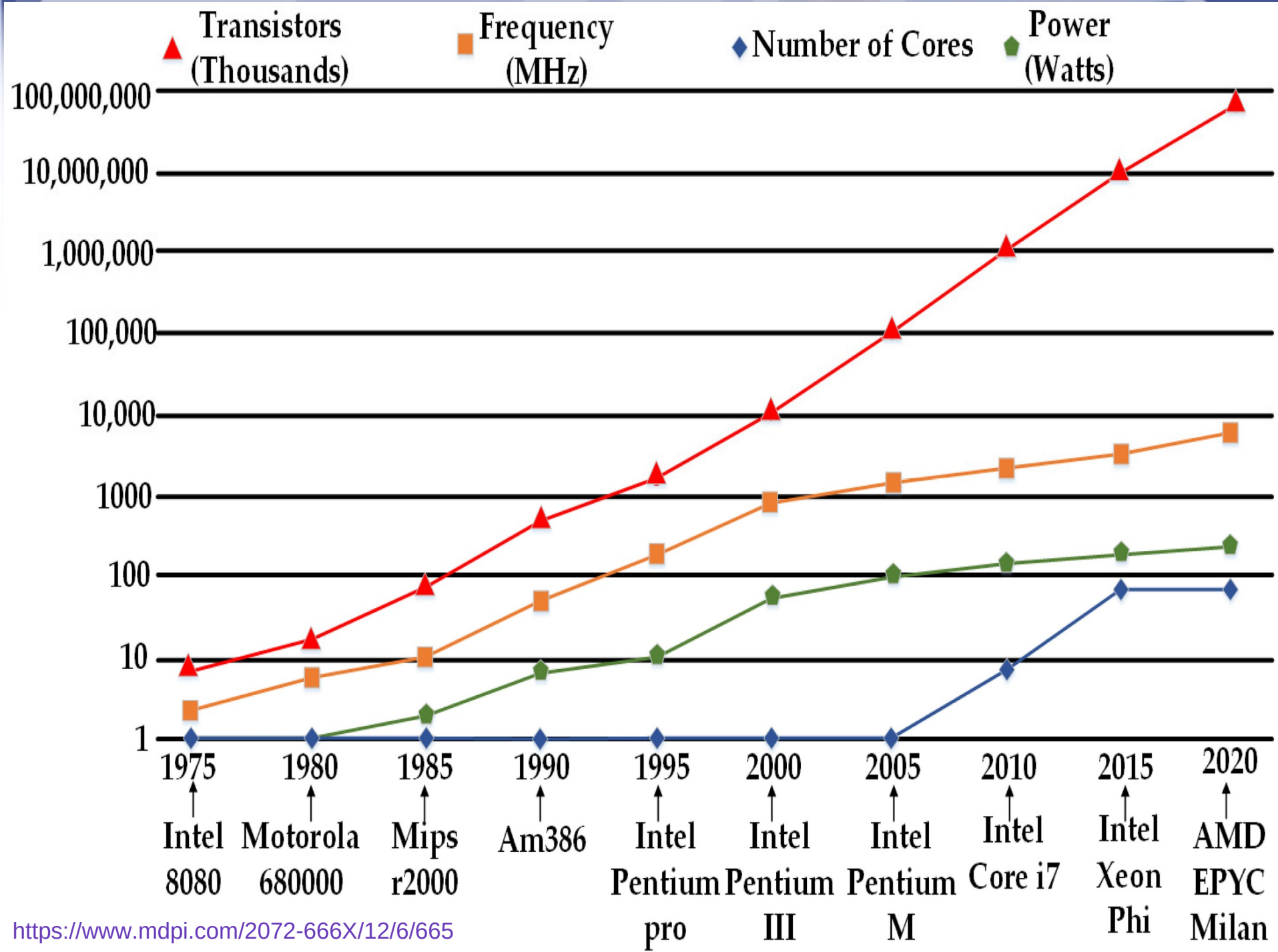
VLIW processors

- IA64
- RISCV

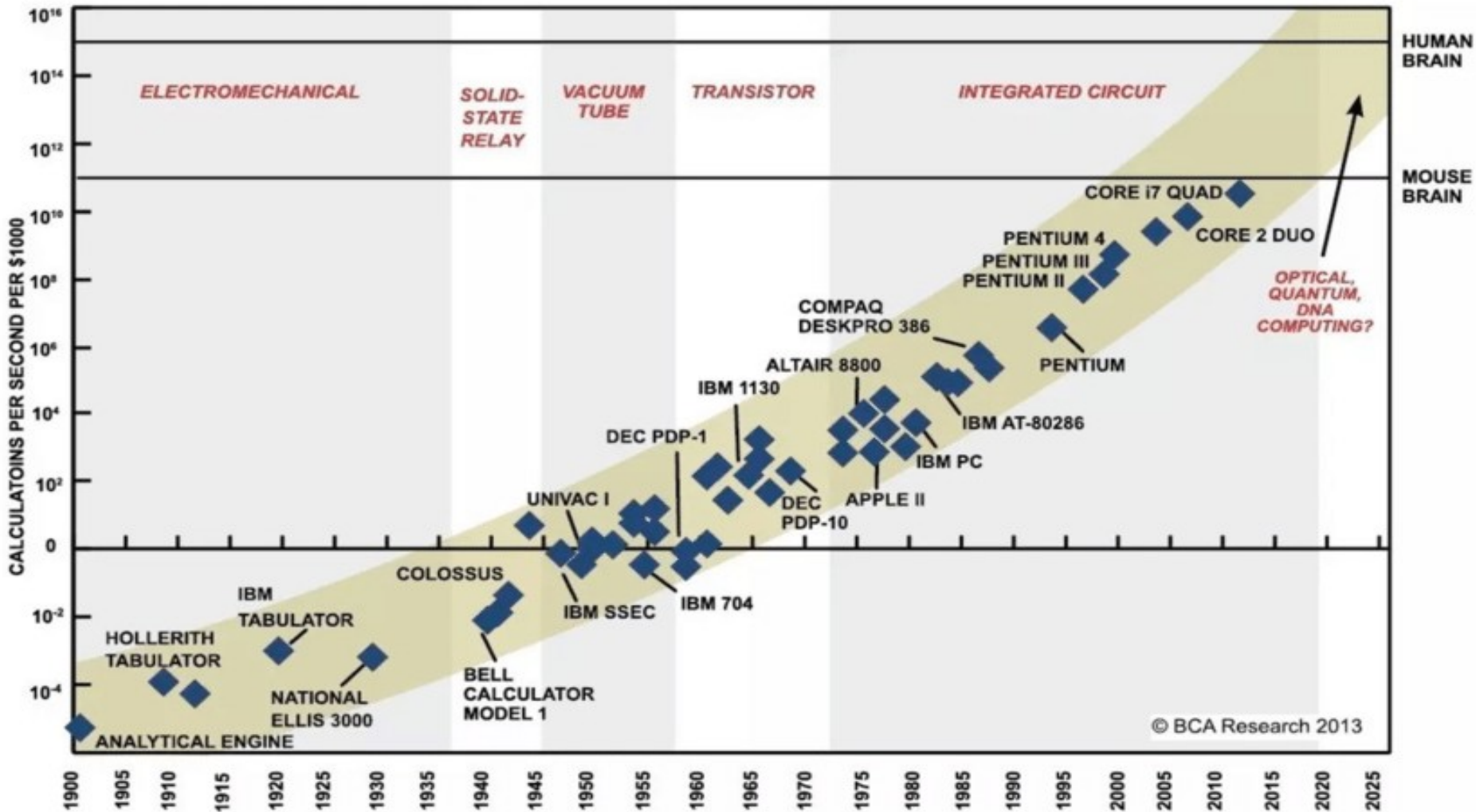
Vector processors

- NEC SX-6
- Cray (Cray X1)

Name	Bit-Width	Year of Invention	Number of Instructions	Clock Speed (MHz/GHz)	Number of Transistors	Instr Per Cycle (IPC)	Operations Per Second (OPS)
Intel 4004/8008	4/8 bits	1971	46	0.074 MHz	2300	1	0.074 M
Intel 8086	16 bits	1978	117	5 MHz	29000	1	5 MIPS
Intel 80386 (386)	32 bits	1985	386	16 MHz	275000	1	16 MIPS
Intel Pentium (P5)	32 bits	1993	~300	60-66 MHz	3.1 million	1-2	60-132 MIPS
Intel Core i7 (Nehalem)	64 bits	2008	~1,000	2.66-3.33 GHz	731 million	2-4	5.32-13.32 GFLOPS
AMD Ryzen 9 5950X (Zen 3)	64 bits	2020	~1,200	3.4-4.9 GHz	10.4 billion	4-6	13.6-29.4 GFLOPS
RISC-V	32/64 bits	2010 (ISA)	47 – Extendable	3 – 5 GHZ	7 Billion	2-Varies	TFLOPS (SoC)



Cost Vs Performance: Electromechanical to ICs



SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, THE VIKING PRESS, 2006. DATAPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

1 Operation / Second = 1 B\$

1B Operation / Second < 1\$

What kept alive the Moors Law

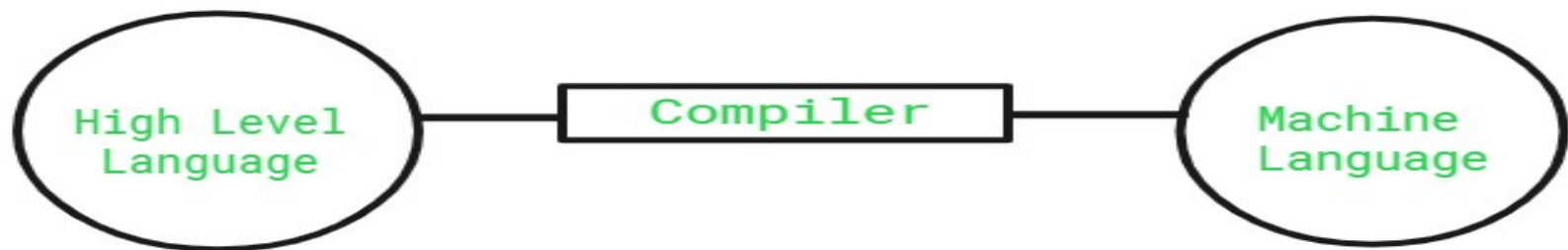
- Technological Innovation
- Market
- Chip Industry
- Digital Data

Birth of Compilers

Ada Lovelace wrote algorithm for **Calculating Bernoulli Numbers** which is often considered the world's first computer program. It consisted of a series of steps and operations that would be performed by the machine to compute these mathematical values.

"COBOL" (Common Business-Oriented Language)
in the late 1950s.

Record-keeping, data validation, and report generation.



Complexity of Processors

Processors have superscalar, long pipelines, and complex internal structures, and they support vector extension units in the CISC RISC architecture.

For high-performance executable programs, modern compilers must also have high performance themselves.

Faster compilers (build tools) are critical for achieving high productivity for large market.

GCC Revolutionized the Software Industry

User controls the Program, FreeSoftware

GCC-1.0: Released by Richard Stallman in 1987.

GCC-2.0: Released in 1992 and supported C++.

GCC-3.0: Released 2001, Developers strong desire for good compilers.

GCC-4.0: Released in 2004

GCC-5.0: Released in 2015 after that each version every year.

GCC-12

Revolution in Computing

Linux is a versatile and widely-used open-source operating system that has revolutionized the world of computing.

Linus Torvalds developed in 1991, Linux has become a cornerstone of modern technology, powering a diverse array of applications across various domains.

Software updates in Linux are easier and faster.

Customization allows users to add or delete a feature as needed.

Reliable Scheduler, Memory Manager and Secure File System

GCC and Linux Revolutionized Software and Computing

Open Source:

Multi-Language Support:

Cross-Platform:

Optimizations:

Standard Compliance:

Modularity:

Diagnostics:

Debugging Support:

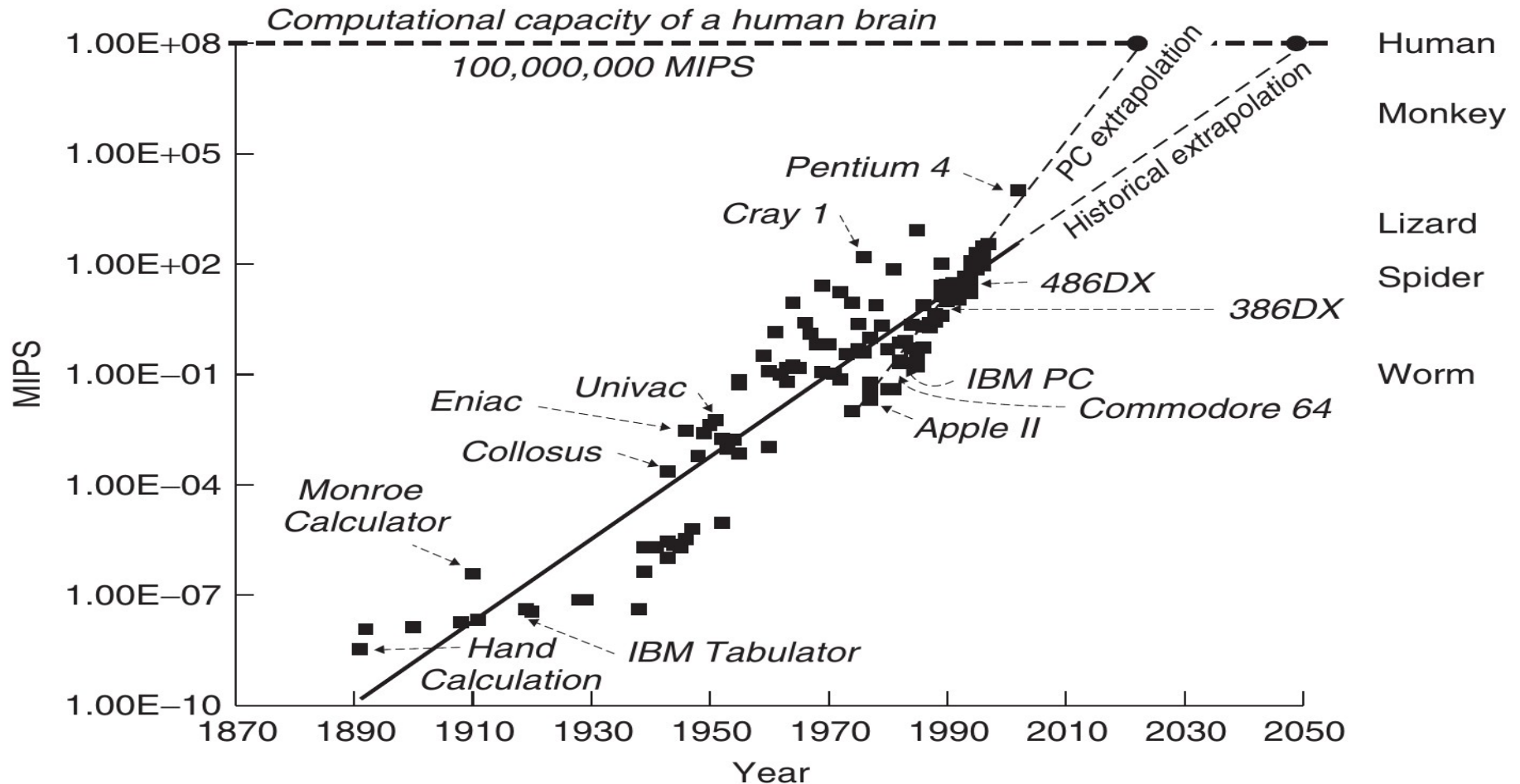
Extensions:

Community and Documentation:

Portability:

Free Software Philosophy:

AI Algorithms and Intellectuality: By Enhancing Computational Capability ?



It is estimated that sometime between the years **2025** and **2050**, a **personal computers** will exceed the calculation power of a human brain.

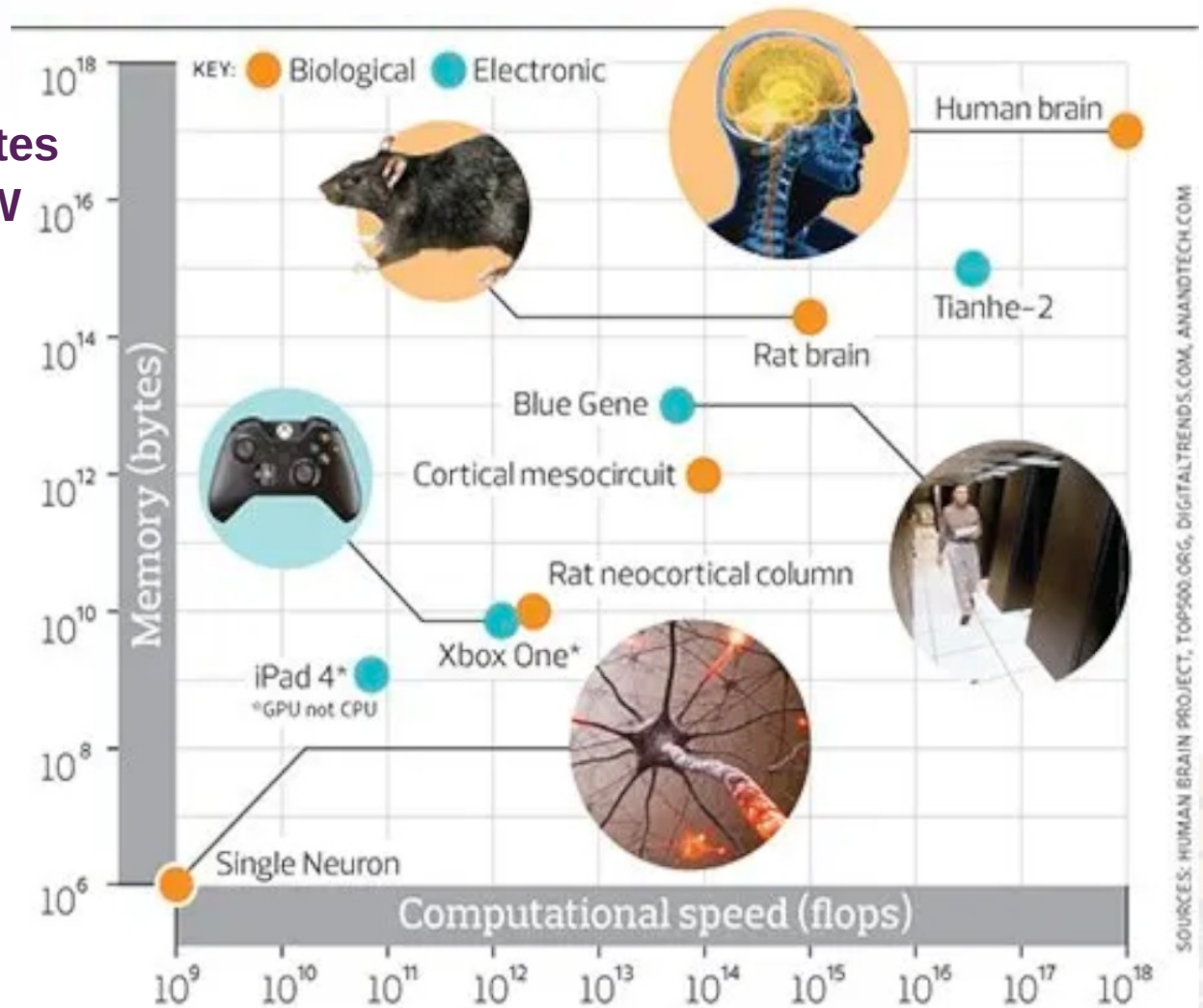
Compute Vs Intellectual Capability

- Human Brain:

- 1 Exa FLOPS

- Memory 100 Peta bytes

- 100 Tera OPS @ 10 W



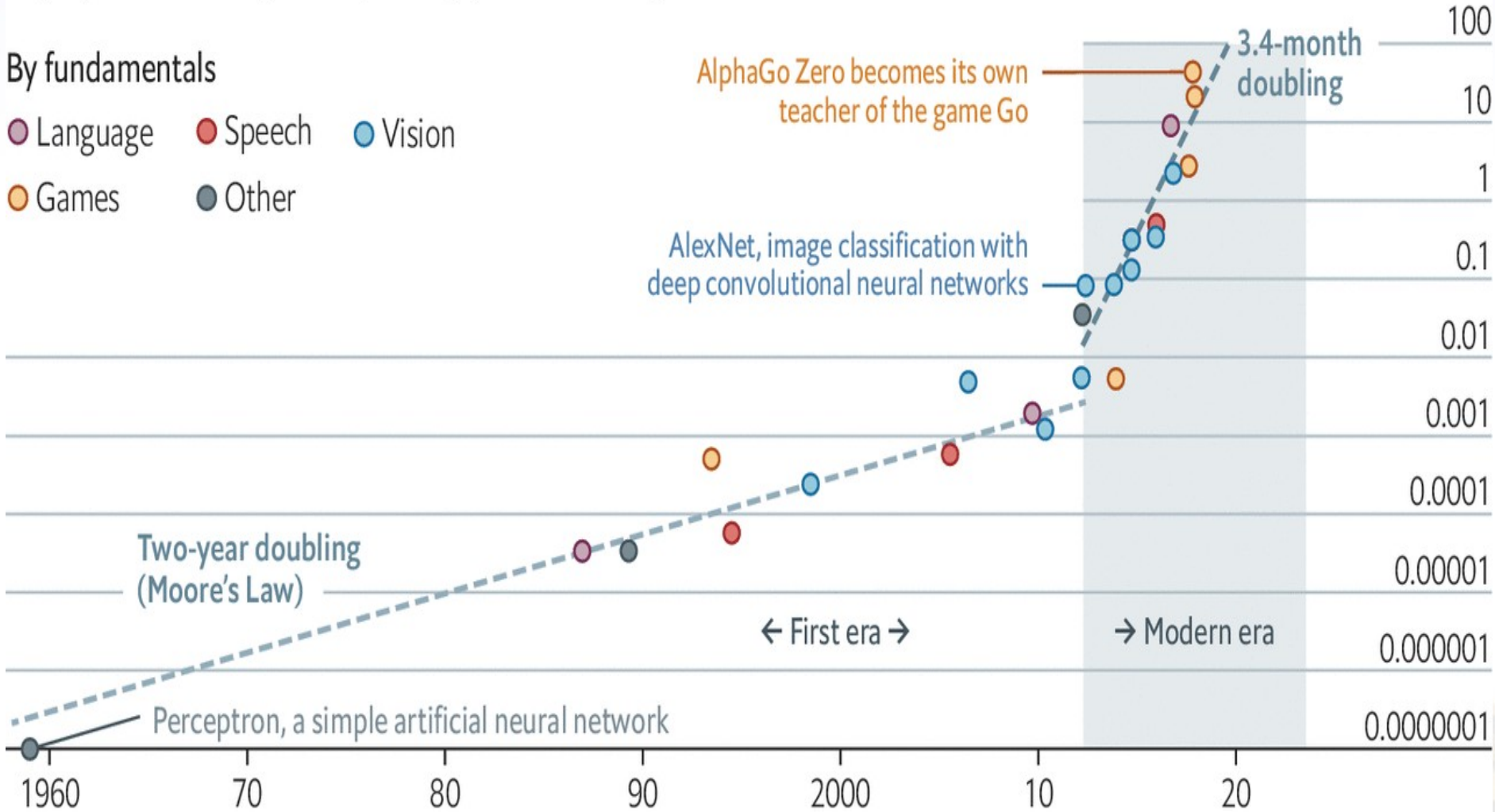
Deep and steep

Computing power used in training AI systems

Days spent calculating at one petaflop per second*, log scale

By fundamentals

- Language
- Speech
- Vision
- Games
- Other

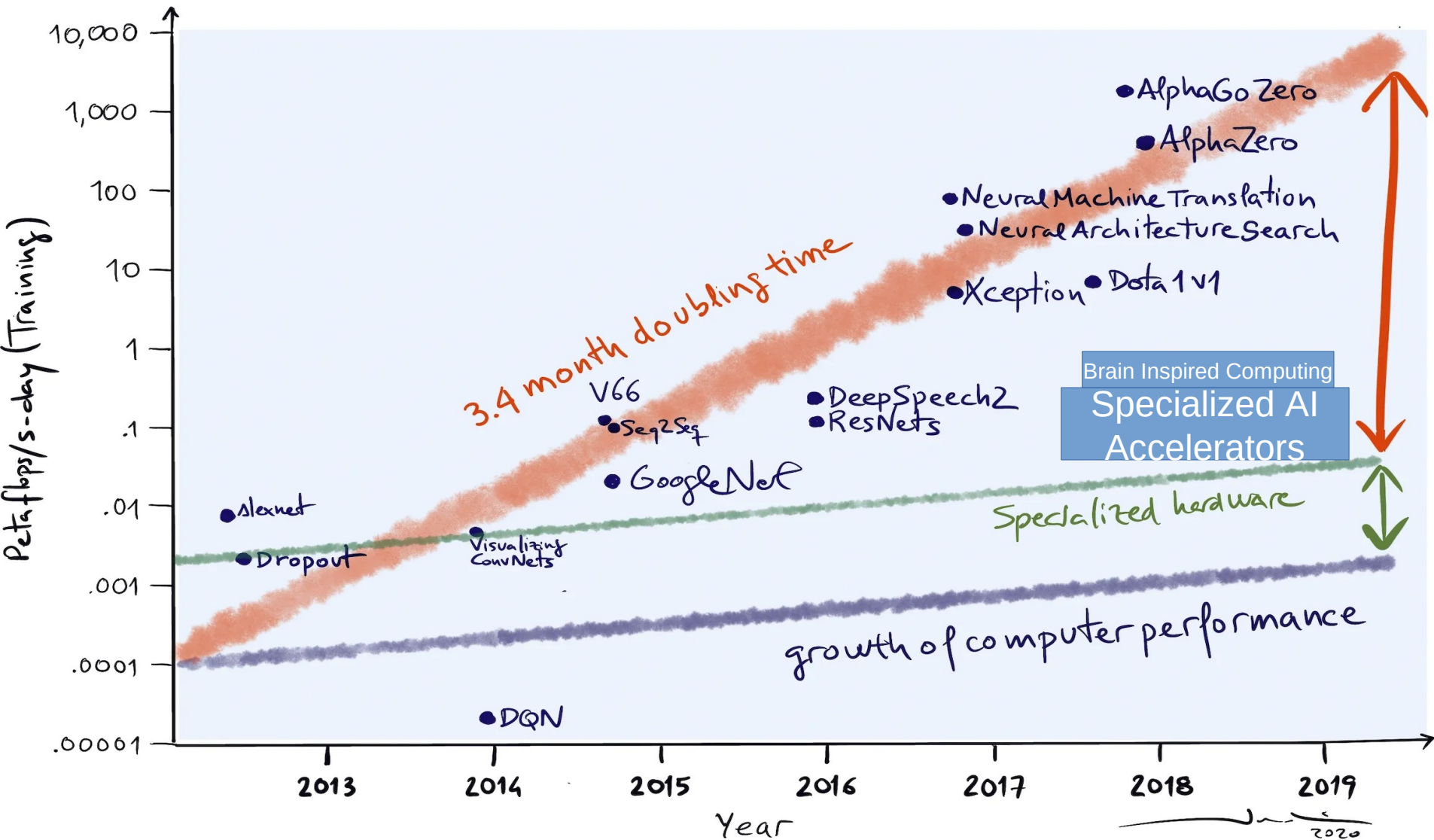


Source: OpenAI

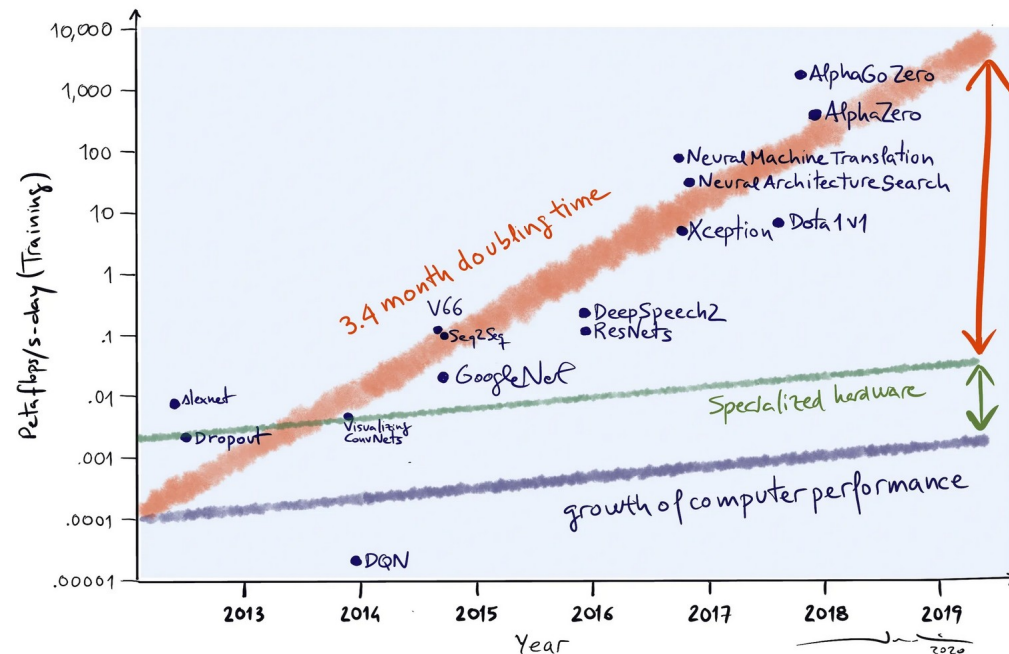
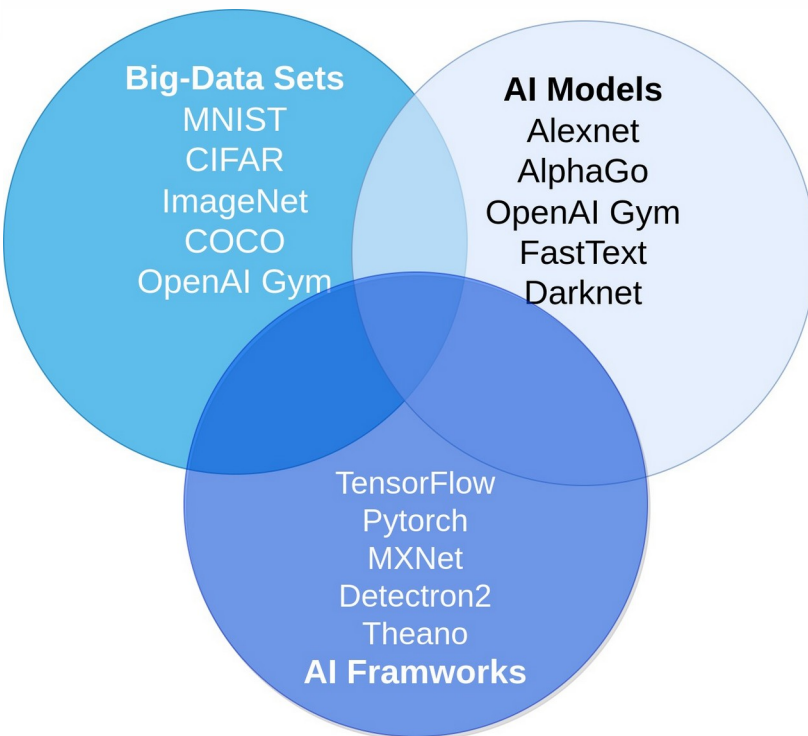
The Economist

*1 petaflop=10¹⁵ calculations

AI and Specialized Accelerators Performance Gap



Open-Source Software: Compilers, Mathematical Algorithms and Data-sets



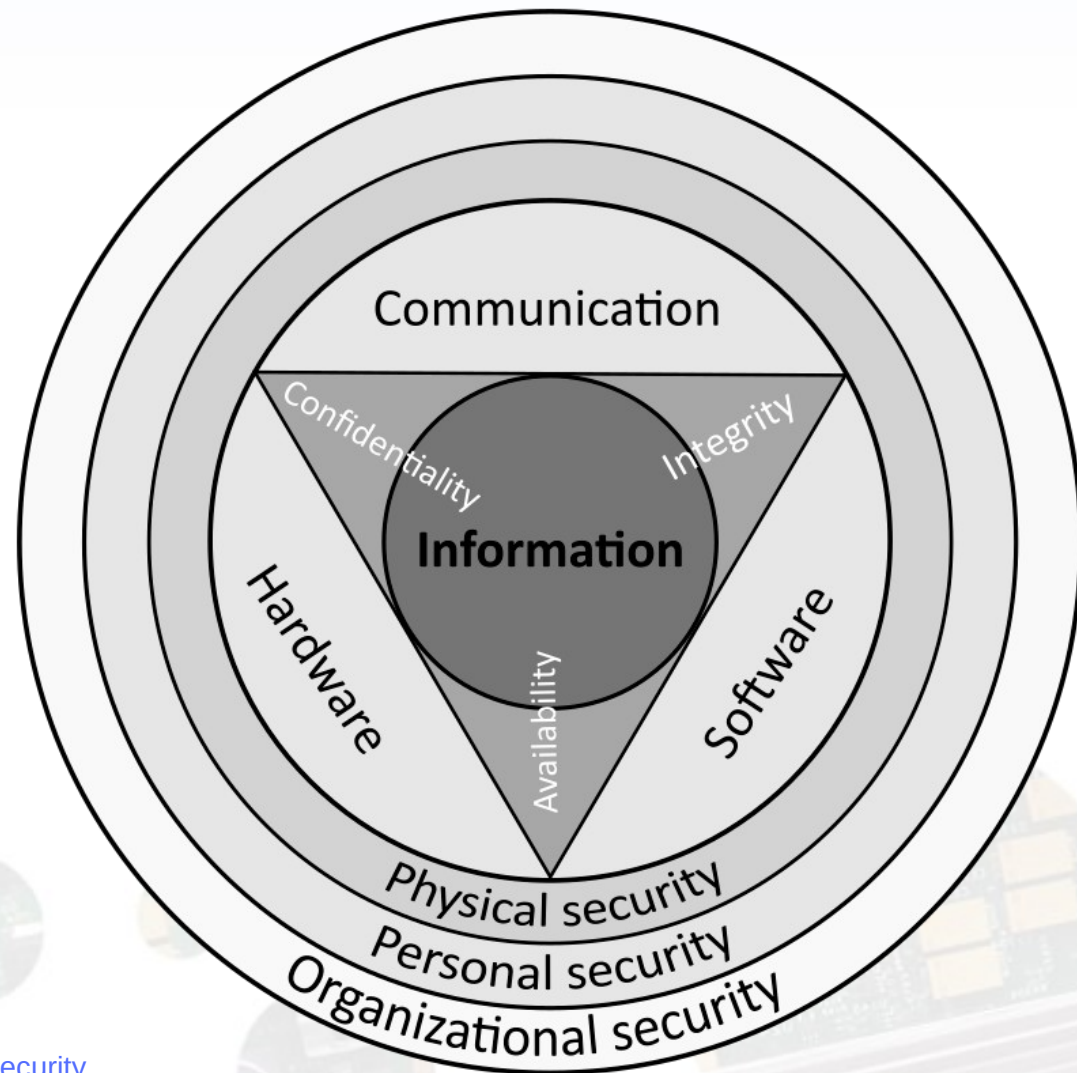
Mathematical Algorithms, Big datasets, and open-source DL framework, play an important role to create “big” algorithms.

Power Wall

Memory Wall

Performance Wall

Security Wall



Open Source: Hardware and ASIC Tools

Democratizing Innovation

Customization and Flexibility

Reducing Costs

Accelerating Development

Community Collaborations

Transparency and Trustworthiness

Reduced Dependence on Proprietary Solutions

Secure Boot and Trusted Execution

RISCV Arch

What is RISC-V: RISC-V is an open, free, and extensible ISA that provides a framework for creating custom processor designs.

Origin: RISC-V was developed at **UC Berkeley**, and it has gained global momentum as an open-source alternative to proprietary ISAs.

Key Principles: RISC-V adheres to key principles, including simplicity, modularity, and scalability, making it suitable for a wide range of applications.

Advantages of RISC-V:

Open Source: RISC-V is open source, which means anyone can access, use, and modify it without licensing fees or restrictions.

Customization: RISC-V is modular, allowing for easy customization of processor designs to meet specific needs.

Diverse Ecosystem: RISC-V has a growing ecosystem of hardware, software, and tools, including compilers, simulators, and development boards.

Current Use Cases and Future

Edge Computing: RISC-V is commonly used in embedded systems, IoT devices, and microcontrollers due to its low-power and flexibility.

Supercomputing: It's also gaining traction in high-performance computing (HPC) and data centers, where custom accelerators are crucial.

Opensource Tool-chains: Ported the compilers and Linux and got other operating systems up and running

Standardization: Expect further standardization of RISC-V ISA extensions, making it easier to develop compatible hardware and software.

Accelerated Adoption: Continued growth in industry adoption, with more companies leveraging RISC-V for their products and services.

Security Enhancements: Focus on security extensions and features to make RISC-V-based systems more secure against emerging threats.

Education and Research: RISC-V will continue to be a valuable educational tool and a platform for cutting-edge research in computer architecture.

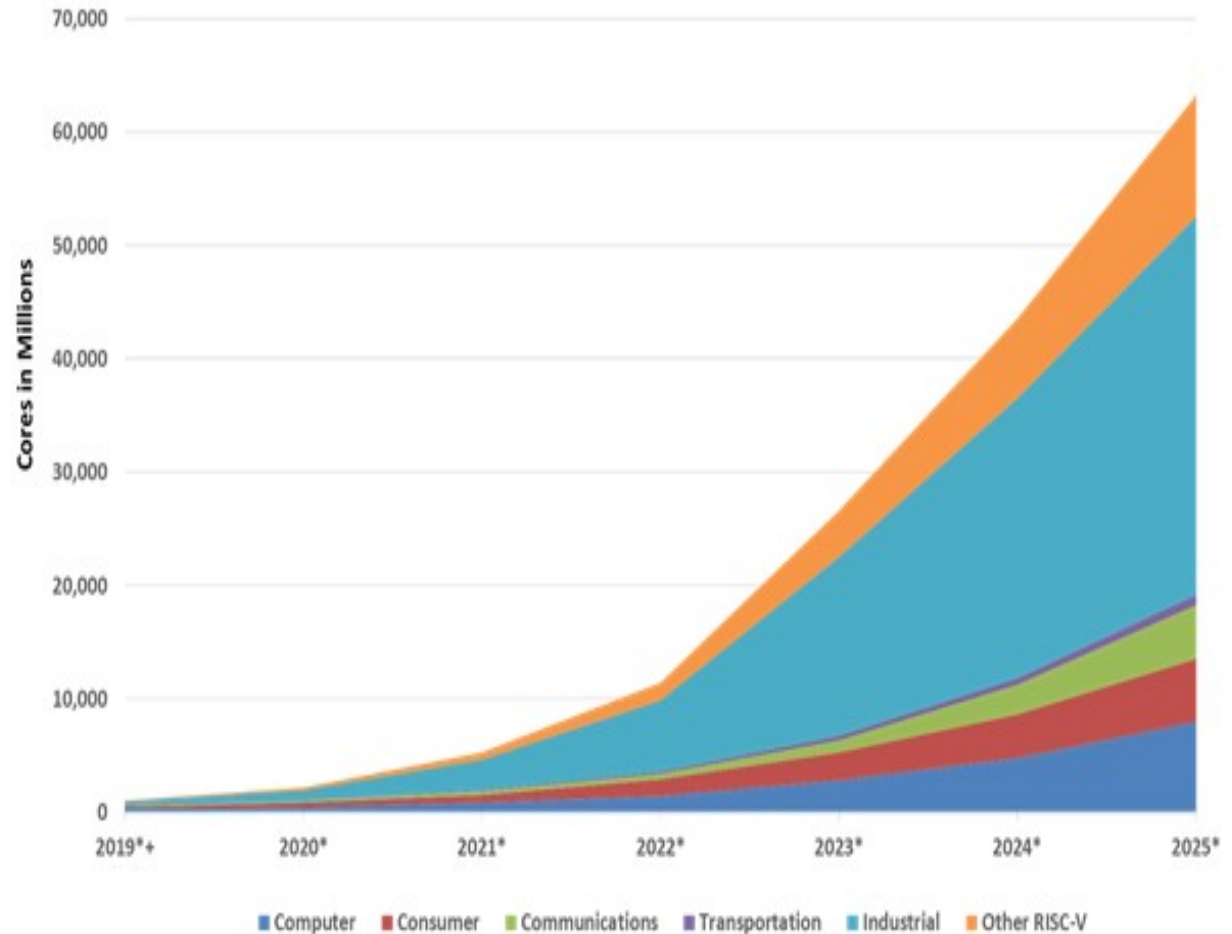
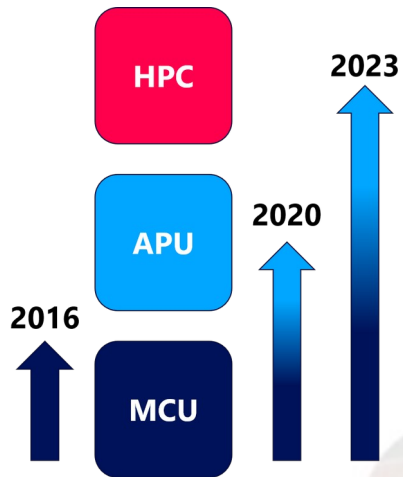
Further Customization: Expect more customized processors and domain-specific architectures tailored to niche applications.

Community Engagement: The RISC-V community will remain active and collaborative, with forums, conferences, and workshops fostering knowledge sharing.

Future of RISC-V



52.22 teraFLOPS
Esperanto Technologies Supercomputer-on-Chip 1

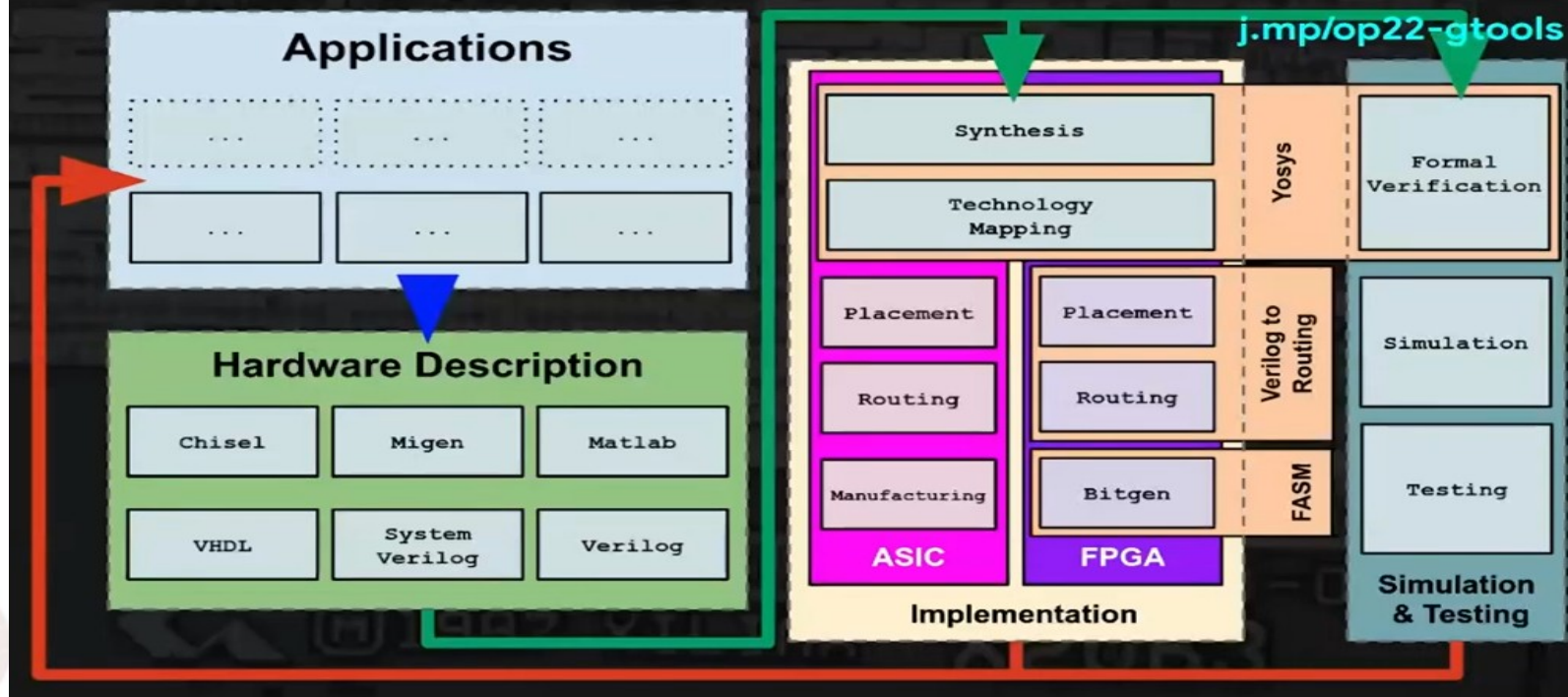


RISC-V shipments predicted to grow strongly. Source: Semico Research Corporation.

4




is now changing hardware development



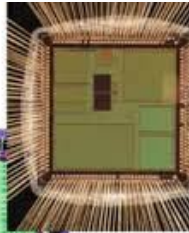
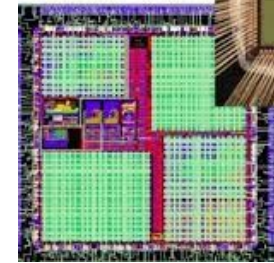
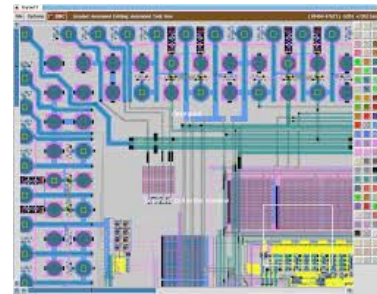
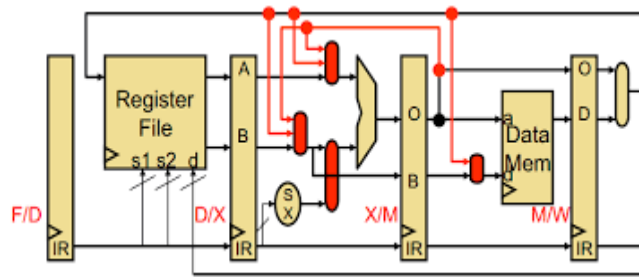
Hardware Design Going to Follow Journey of Software Design

	Software	Hardware
High-Level Languages	Python, Ruby, R, Javascript, Julia	Chisel, PyMTL, PyRTL, Myhdl, JHDL, Clash, Calyx, Dfiant
Libraries	C++ Std, Python Std Libs	Basejump
Tool Chains	GCC, LLVM, CPython, MRI, PyPy, V8	Icarus Verilog, Verilator, Qflow, Yosys, Timberwolf, Qrouter, Magic, Klayout, Ngspice
Standards	POSIX	RISC-V ISA, ROCC, Tilelink
Systems	Linux, Apache, Mysql, Memcached	RocketChip, Pulp/Ariane, OpenPiton, ChipYard, BOOM, FabScalar, MIAOW, Nyuzi
Methodologies	Agile Software Design	Agile Hardware Design
Cloud	IaaS, Elastic Computing	IaaS, Elastic Cad

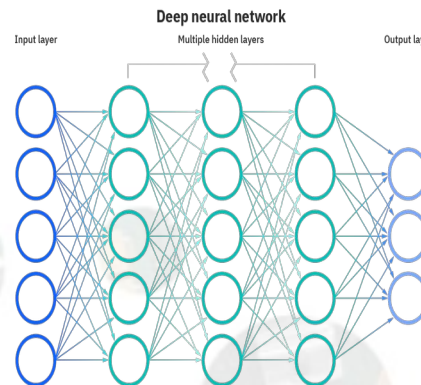
- 
- Mankind Progress
 - Age of Big Data and AI
 - Micro-electronics! Revolutionizing the World
 - **Namal Chip Design and HPC Facility**

Microelectronics Solutions for AI Compute Capability

- OpenSource Full-Stack Ecosystem for RISC-V Processor System



- Supercomputing for AI and BigData Applications



OpenSource Full-Stack Ecosystem for RISC-V Processor Architecture

- **Hardware Architecture**

- Low Performance and Low Cost Digital System
- Uni/Multi Core System on a Chip

- **Single Board Computer**

- Hardware Software Co-Design
- High Performance Computing

- **Intelligent and Real-time Applications**

- Industrial Automation
- Machine Learning

Applications Domain

Embedded
Systems

Biomedical
Engineering

Internet of
Things

Agri-tech

Edge Computing

Industrial Automation

Hardware Stack

Internal Bus
Interconnectoin
AMBA, WISHBONE

Accelerators
DSPs/FPGAs

Processors
RISC-V
RISC/MIPS

Memories and
Disks

High-Speed
Ethernet,
Infiniband

Hardware Development Approaches

Switch Level

Gate Level

Rigester
Transfer Level

Behaviroal
Level

Layout

VHDL

Verilog

C++

Data Flow

COTS ASIC

Design and Synthesis Tools

QFlow
OpenRoad

Chipyard

Yosys

LegUp

COTS
Vivado, Cadence etc.

Magic, KLayout

RocketChip

VTR

Chisel

IVerilog
Modelsim

Applicatoin Development Stack

Freedom Studio
PlatformIO

Linux, RTOS

QEMU, Spike

Python, Cuda

Verilator, IVerilog
GTKWave

TensorFlowLite

GCC, EdgeX

riscv-opcodes
riscv-pk

Apache Cassandra

SQLite

riscv-tests

Supercomputing Platform for AI and BigData Applications

- **Bare-Metal and Containerized Cluster Infrastructure:**
 - Distributed Hardware Interfacing, Network Configuration and Distributed Computing Software Deployment
- **Data Center and Cloud Infrastructure:**
 - Storage systems, networking equipment, and software configuration
- **AI Applications for Scientific and Engineering Problems**
 - Distributed AI applications for multi-node bare-metal system
- **HPC Application Parallel Programming**
 - Heterogeneous multi-node parallel processing using parallel programming models

Applications Services

Data Sciences

Health Science

Social Sciences

Agriculture

High Performance
Computing

Modeling and
Simulation

Web
(IoT, VLSI Design)

Development Frameworks and Libraries

Interactive

GCC

Python

OpenMP

MPI

CUDA

OpenACC

OpenCL

TensorFlow

PowerAI

Horovod

DeepSpeed

Hadoop

Spark

Distributed System & Software Stack

OpenHPC, ROCKS

OpenShift, xCAT
Nutanix Acropolis

Open-Stack
Kubernetes

Linux Kernel: OpenPBS, PBS-Pro, SLURM, Ganglia, Open vSwitch, warewolf, Lustre, BeeGFS, Ceph, Mellanox OFED, IPoIB, OpenEth, Network Information Service, ACPI

Rolls, Singularity Image, Docker, Container

Hardware System

Intelligent RACK
infrastructure
PDU, PMS

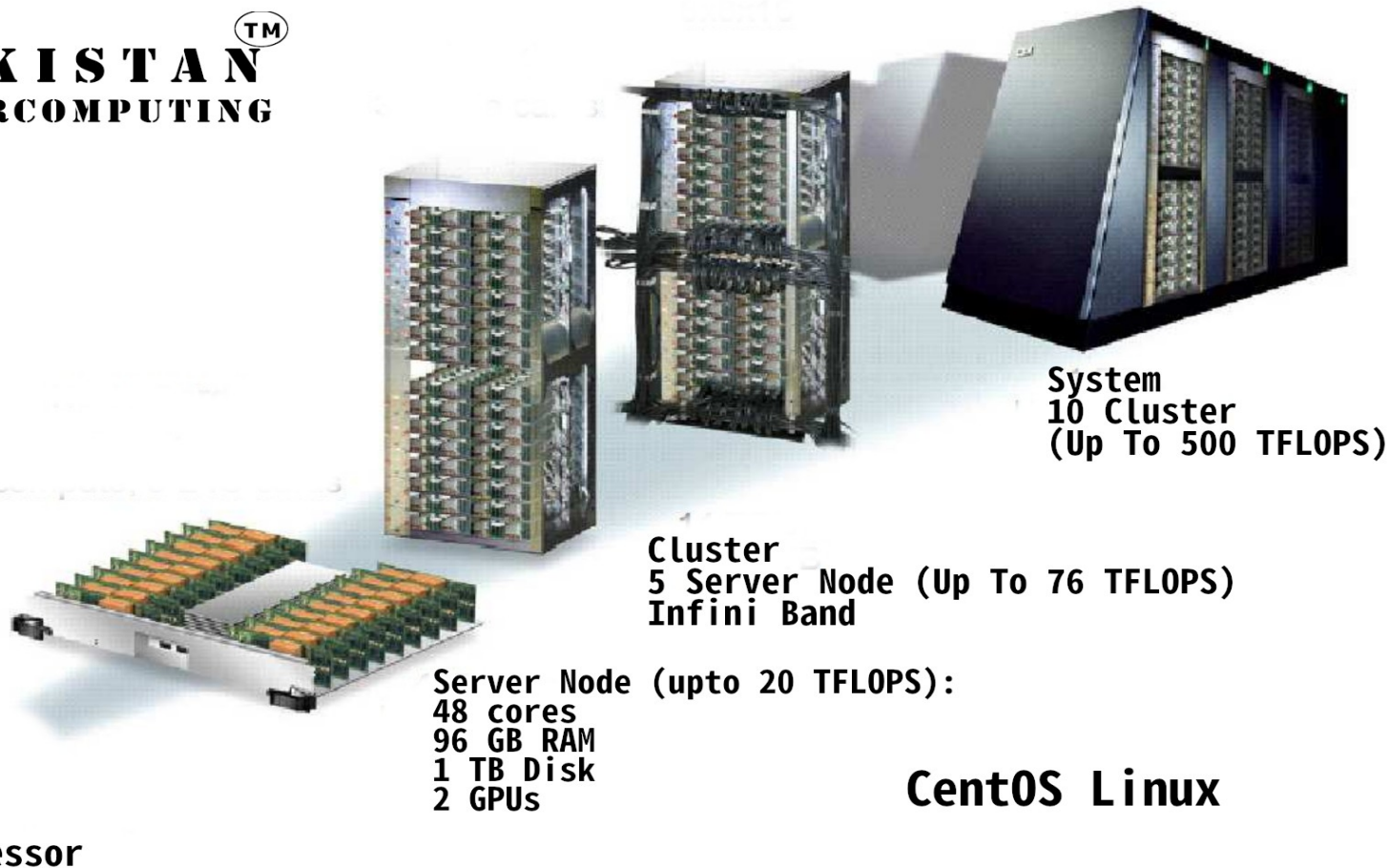
Accelerators
GPU/TPU/FPGAs

Multi-core
CISC/SuperScalar

SAN/NAS,
SSDs/NVMe

High-Speed Ethernet,
Infiniband

Developing Supercomputing for AI



CentOS Linux



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación



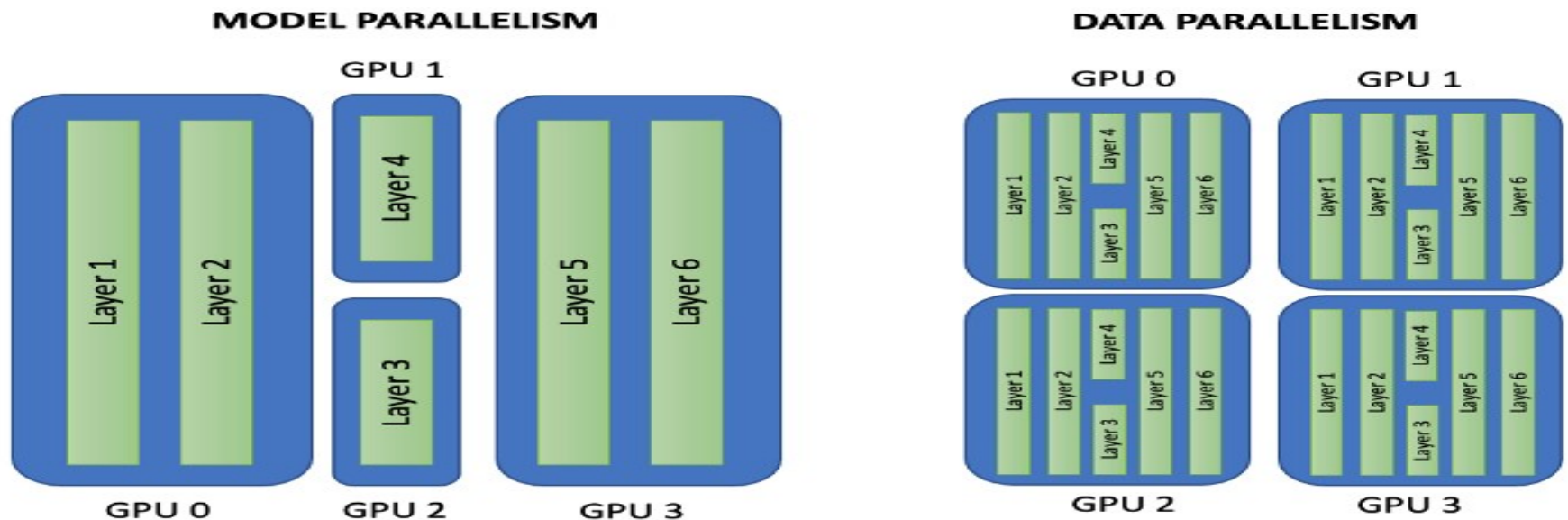
AI Model Parallelism

- **Model Parallelism**

Different layers of the network distributed across different devices

- **Data Parallelism**

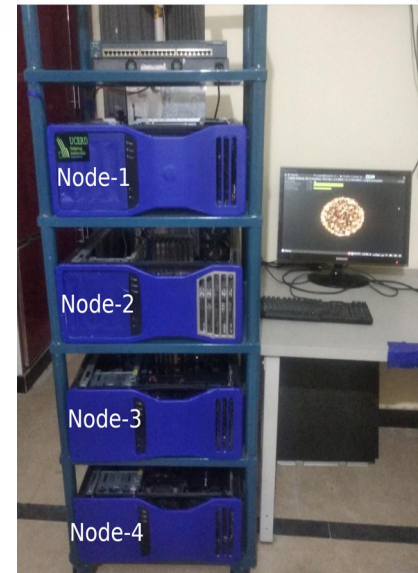
Same model in every one of the GPUs, each processing a separate piece of the data, a separate portion of the mini-batch.



Visit us

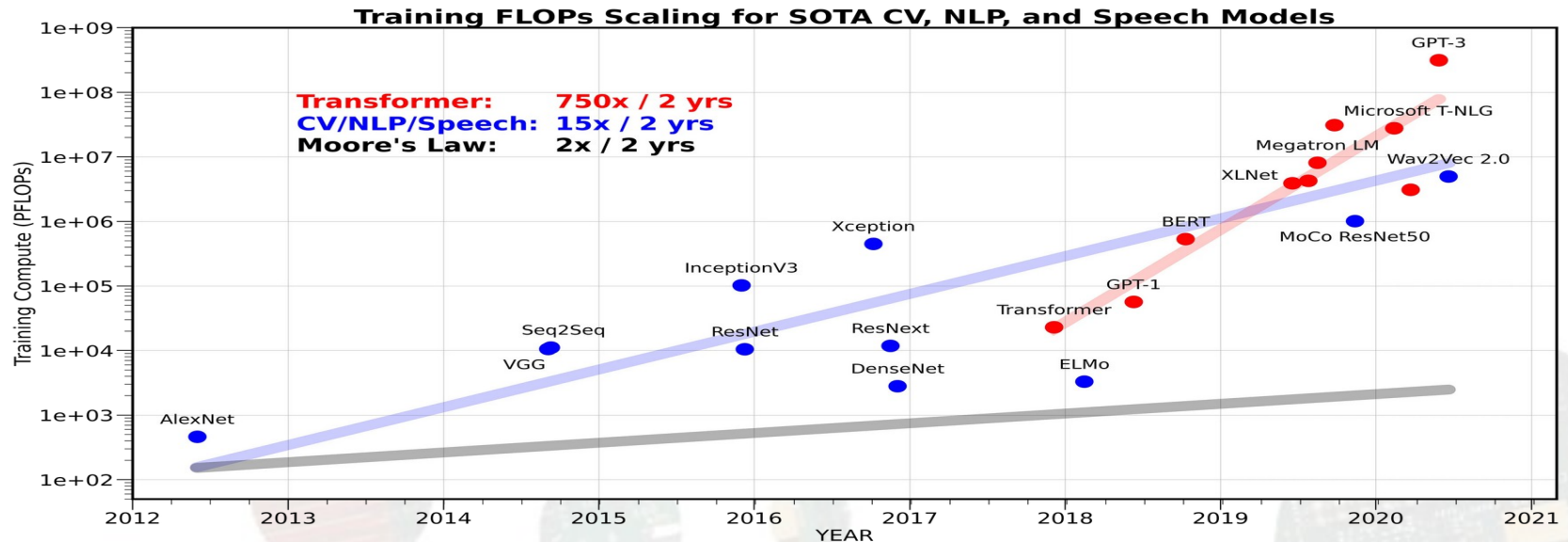
`ssh username@10.0.0.153`

`ssh username@119.156.30.83`



Conclusion

- World Data Size = 130 Zettabytes, doubling every 18 months.
- To handle big-data AI algorithm are the only solution.
- The computational demands of AI algorithms are experiencing exponential growth. (ExaFLOPS/Day)
- Micro-Electronics is the only solution to store big-data and process the AI.



AI is the need of the day and is definitely penetrate society, like electricity

Micro-electronics is the only solution to Handle AI problems.

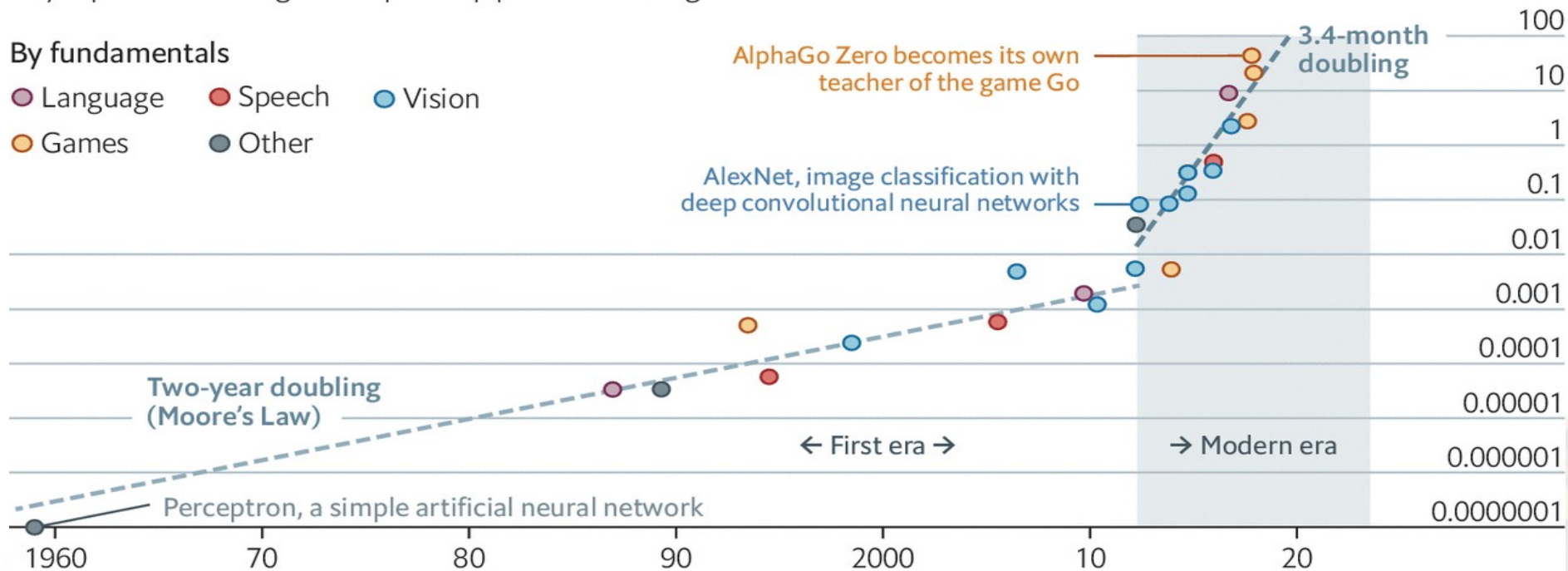
Deep and steep

Computing power used in training AI systems

Days spent calculating at one petaflop per second*, log scale

By fundamentals

Language Speech Vision
Games Other



Source: OpenAI

The Economist

*1 petaflop=10¹⁵ calculations



Center of Excellence:

Supercomputing for

Free Open Source Software Chip Design Stacks and Open Source Processor: Revolutionizing the World

by: Tassadaq Hussain

**Professor Department of Electrical Engineering
Namal University Mianwali**

Collaborations:

Barcelona Supercomputing Center, Spain

European Network on High Performance and Embedded Architecture and Compilation

Pakistan Supercomputing Center