



Center of Excellence:

Supercomputing for

Microprocessor based Systems: Revolutionizing the World

by: **Tassadaq Hussain**

Director Centre for AI and BigData

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

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.



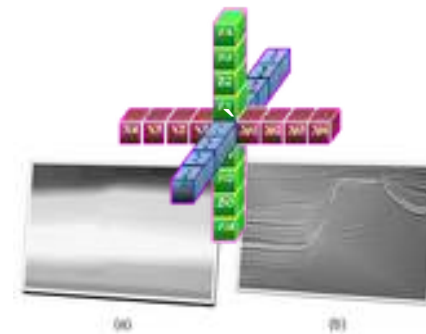
Int'l Projects

- 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**)

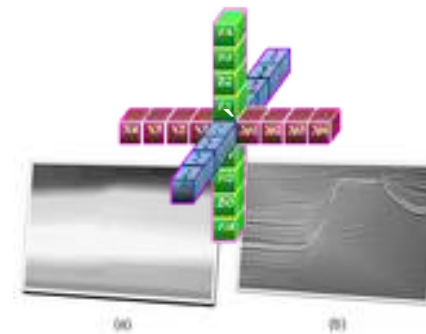


Int'l Projects

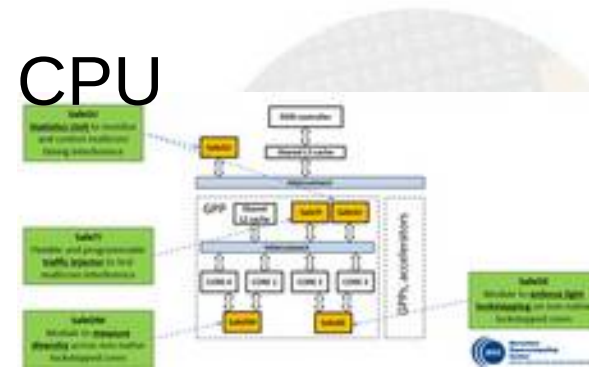
- Design Ultra Low Cost Display Camera Interface for Mobile Baseband XGold Chip (**Infineon Technologies, 200 million single chip**)




- Implementation of Reverse Time Migration on FPGAs (**BSC-REPSOL, PLDA Italia, Cambridge Science Park**)

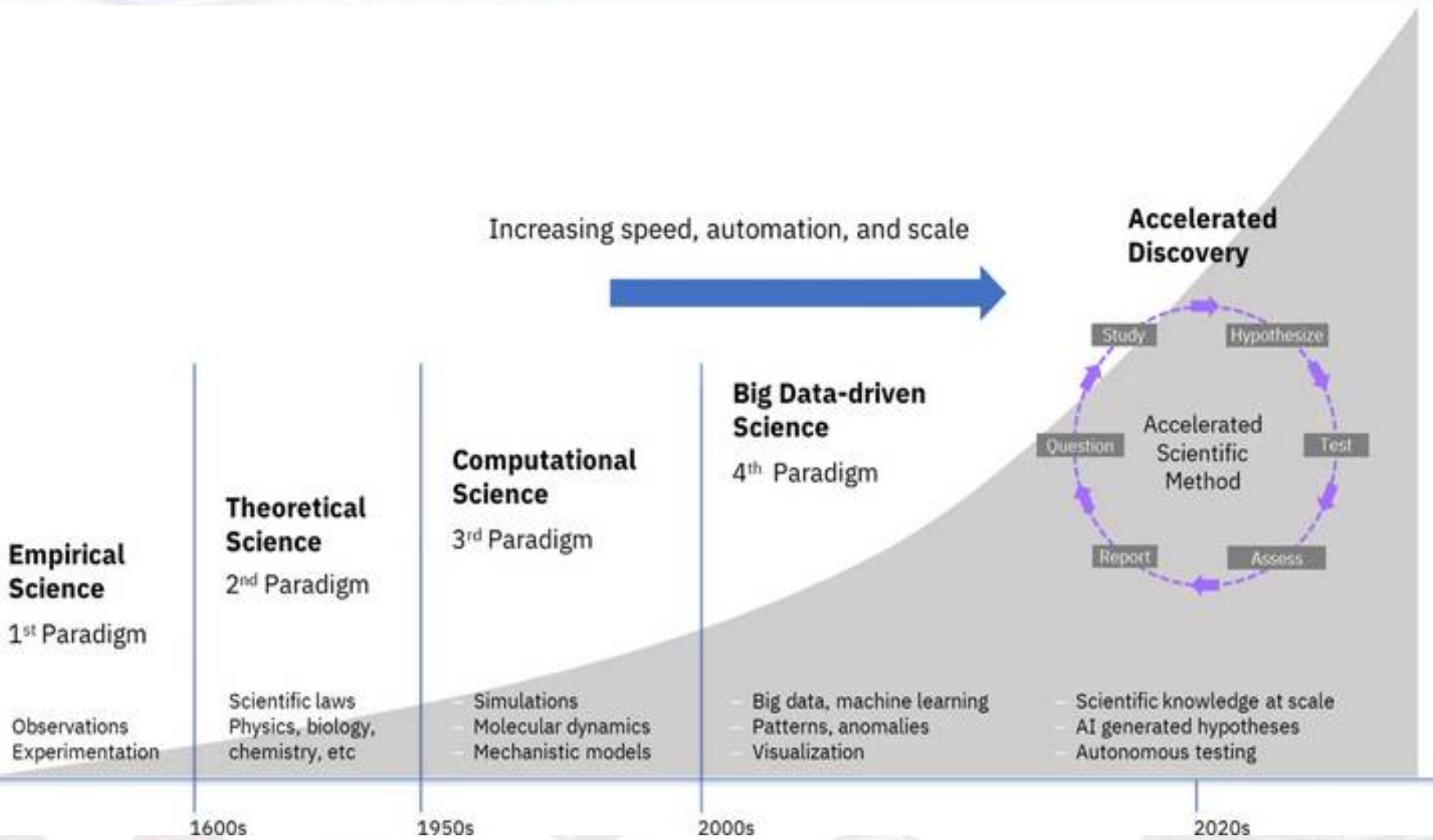


- Open source European full-stack ecosystem based on a new RISC-V CPU (**Barcelona Supercomputing Center**)

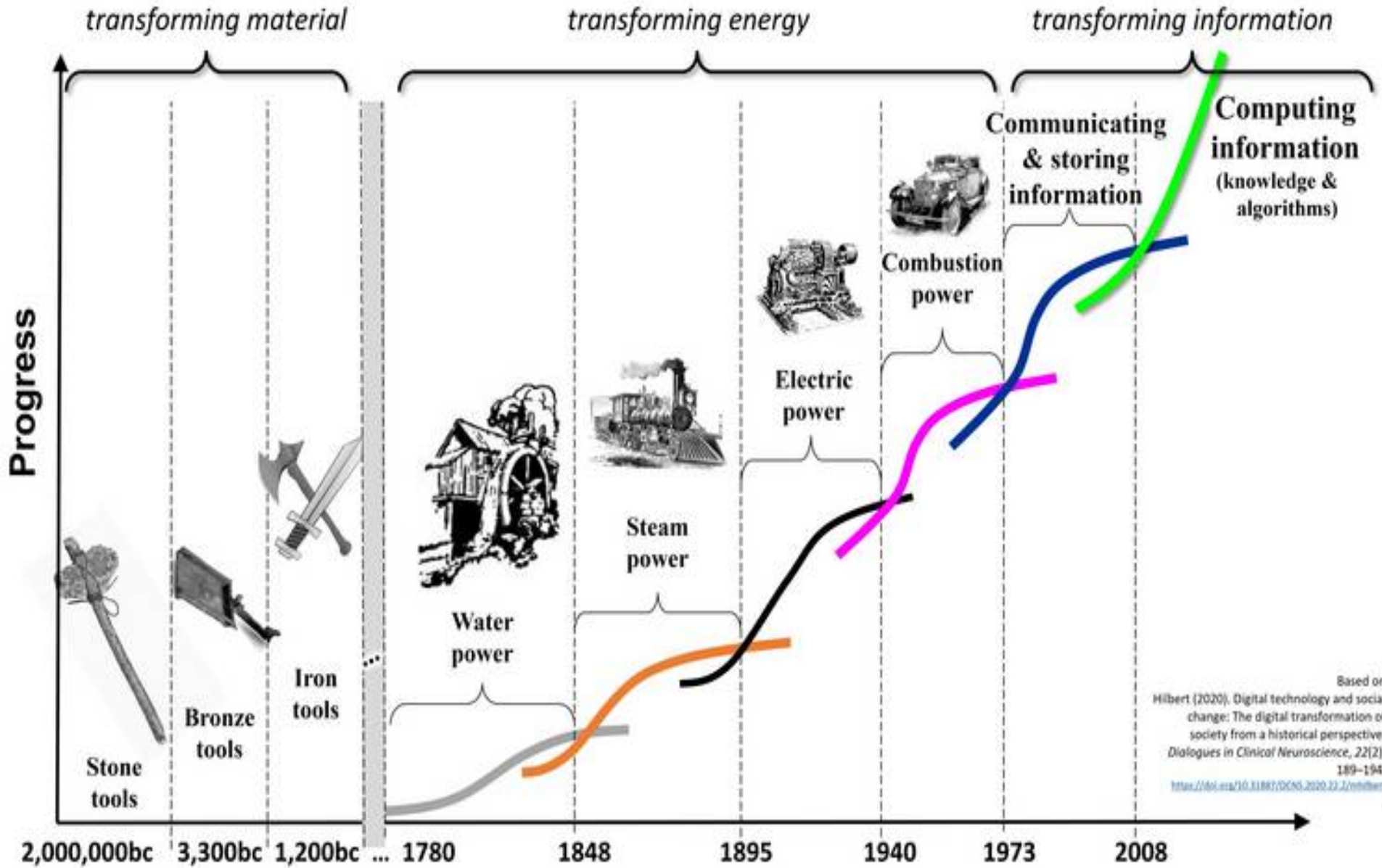


- 
- **Mankind Progress and Industrial Revolution**
 - Age of Big Data and AI
 - Micro-electronics! Revolutionizing the World
 - Namal Centre for AI and BigData

From Age of Empirical Science to Data-Science



Mankind Progress



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



Democratization in Microelectronics

- Microprocessors have revolutionized automation.
- GCC has **revolutionized** the software industry.
- Linux has **revolutionized** computing industry.
- Arduino has **revolutionized** embedded computing.
- AI **revolutionizing** the computing intellectuality.
- RISC-V is **revolutionizing** the Secure Computing.

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)

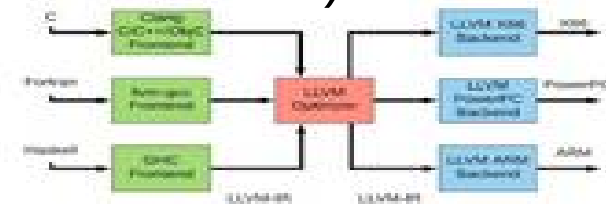
Under Hundred of billion dollars industry



Tier 3:

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

Over Hundreds of Billions of dollars industry




Tier 4:

Hardware Development (Semiconductor, etc)

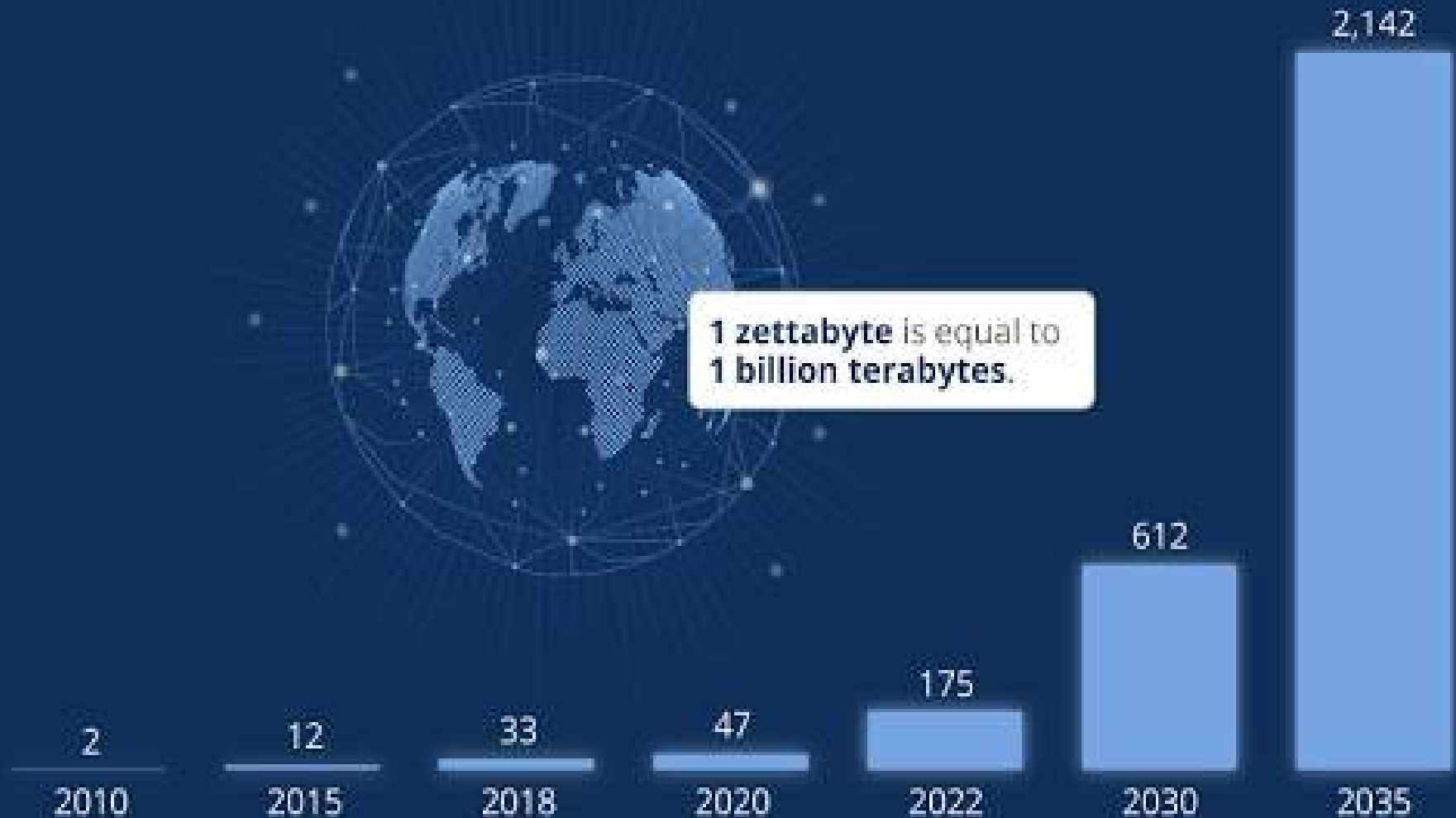
Hundreds of billions to over a trillion dollars Industry



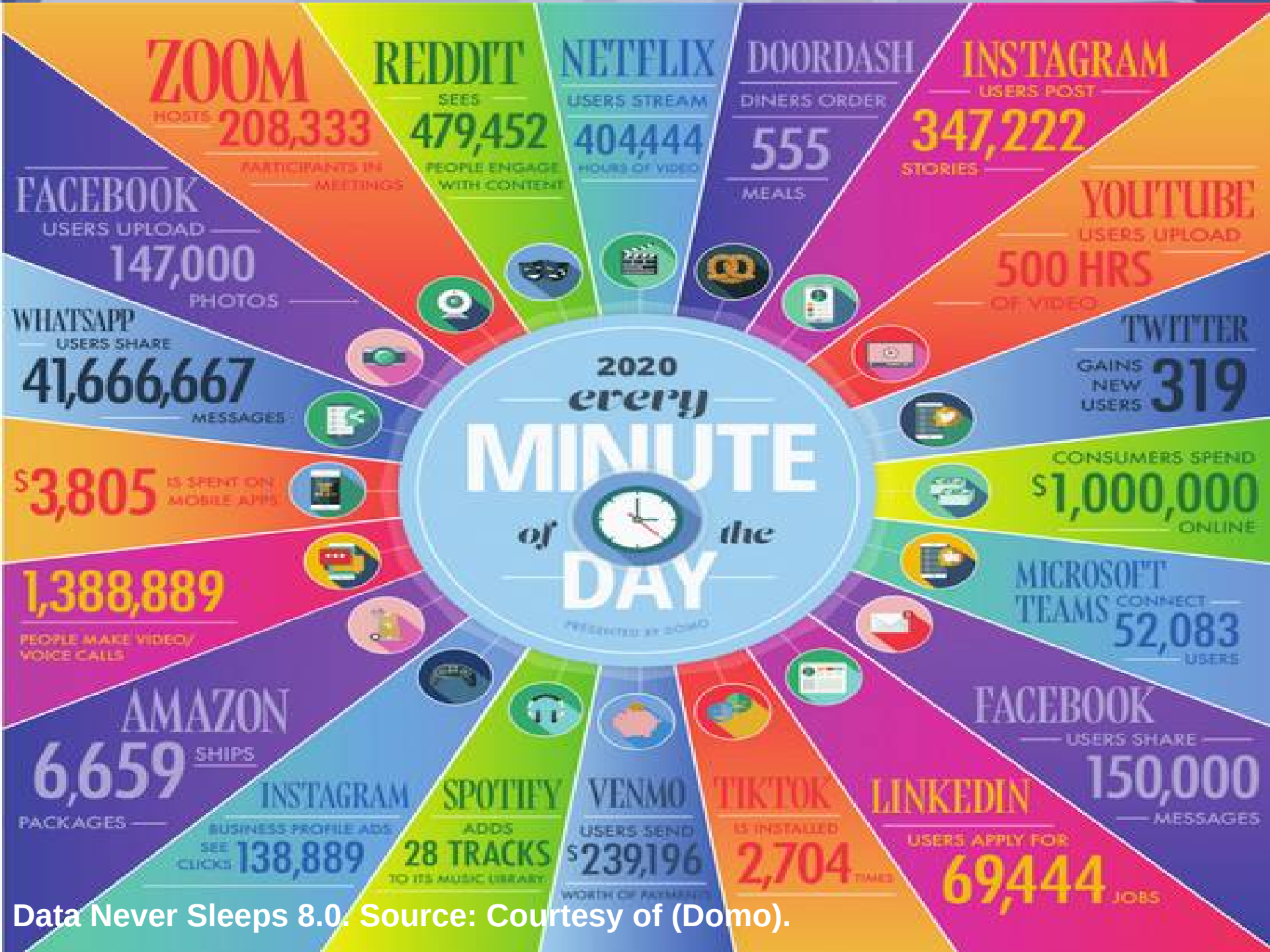
- 
- Mankind Progress and Industrial Revolution
 - **Age of Big Data and AI**
 - Micro-electronics! Revolutionizing the World
 - Namal Chip Design and HPC Facility

Global Data Creation is About to Explode

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

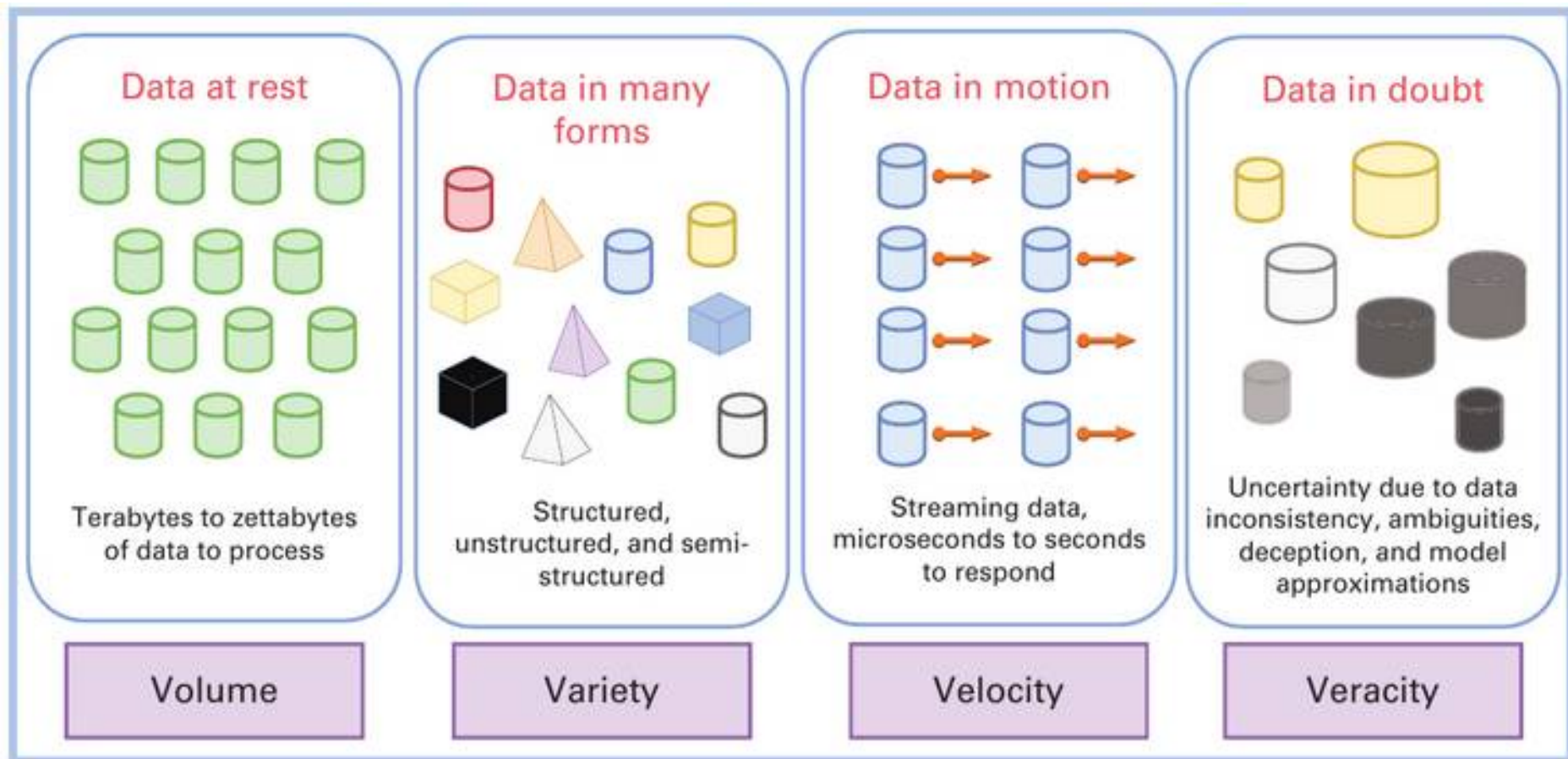


57.76 US\$

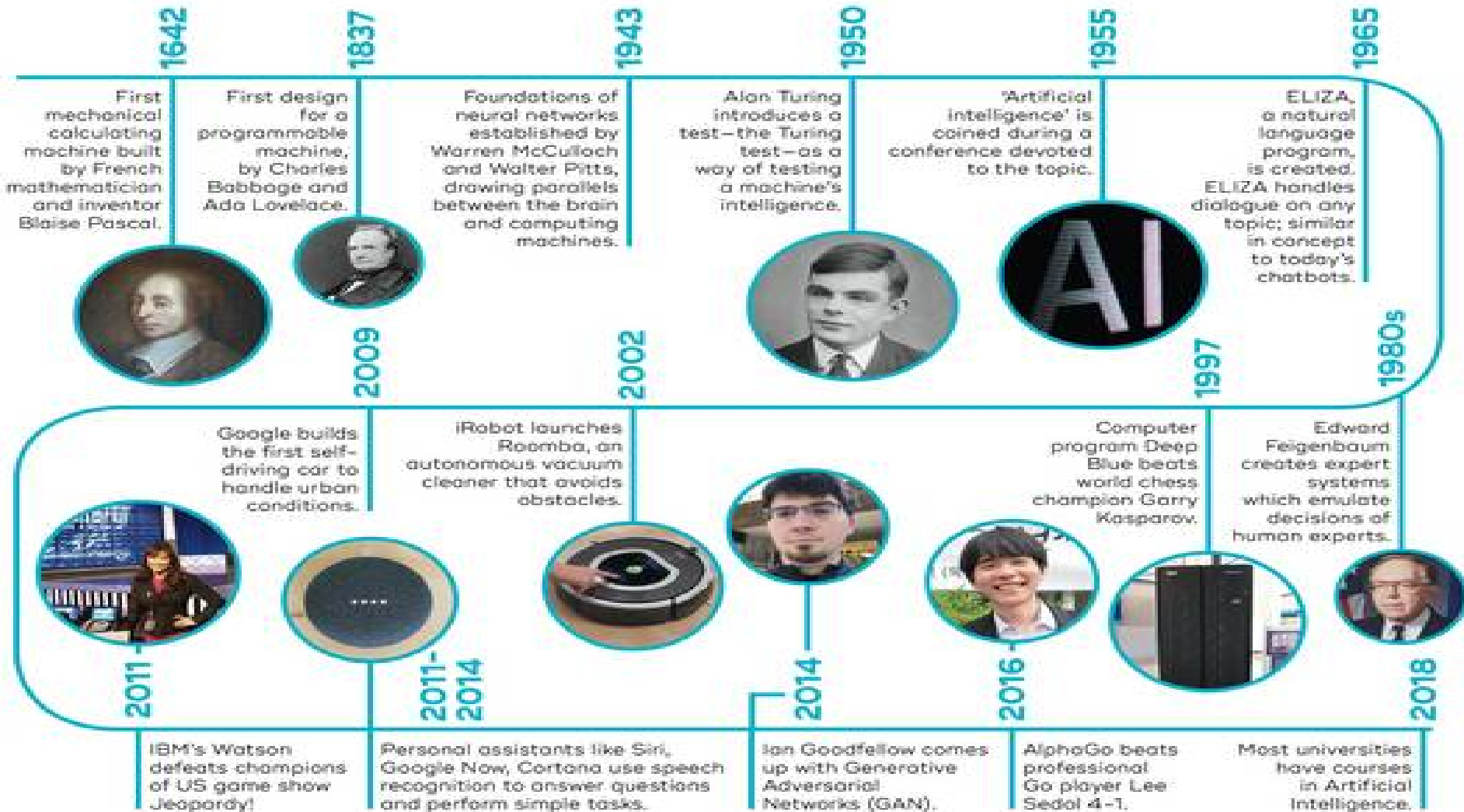


Data Never Sleeps 8.0. Source: Courtesy of (Domo).

Types of Data and its Challenges

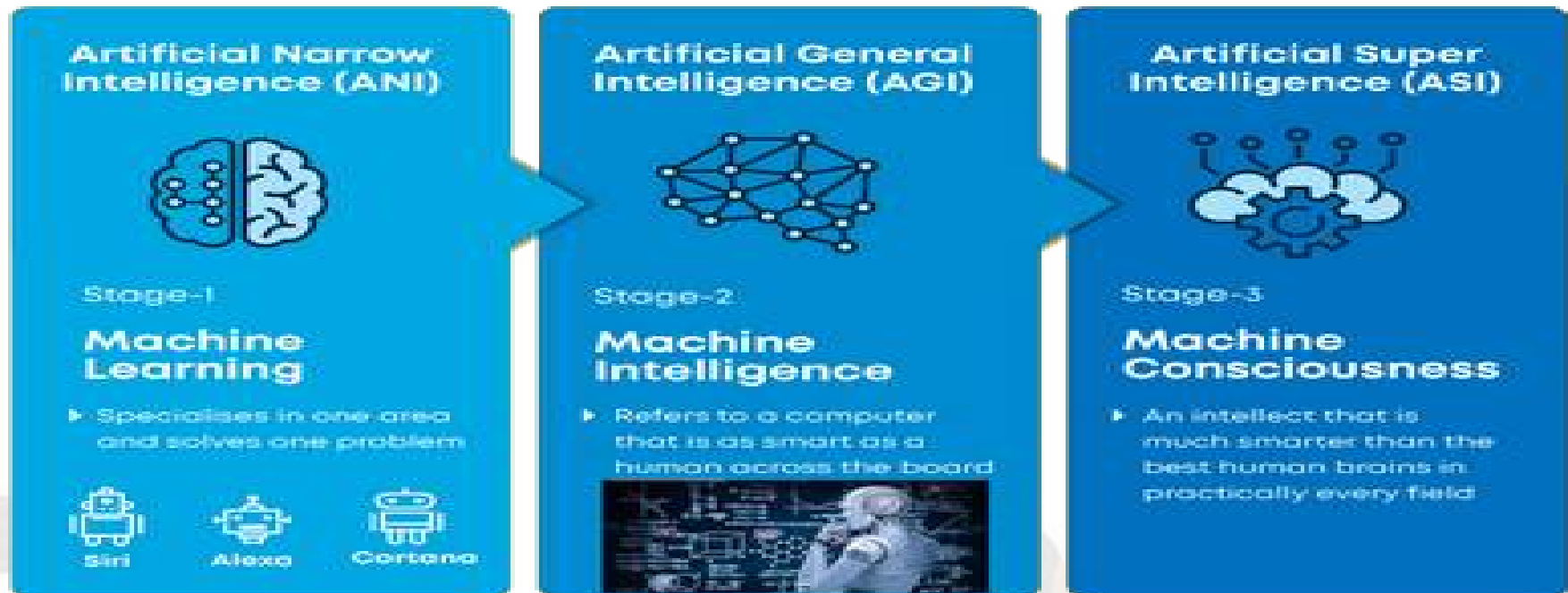


AI: The Only Solution for BigData



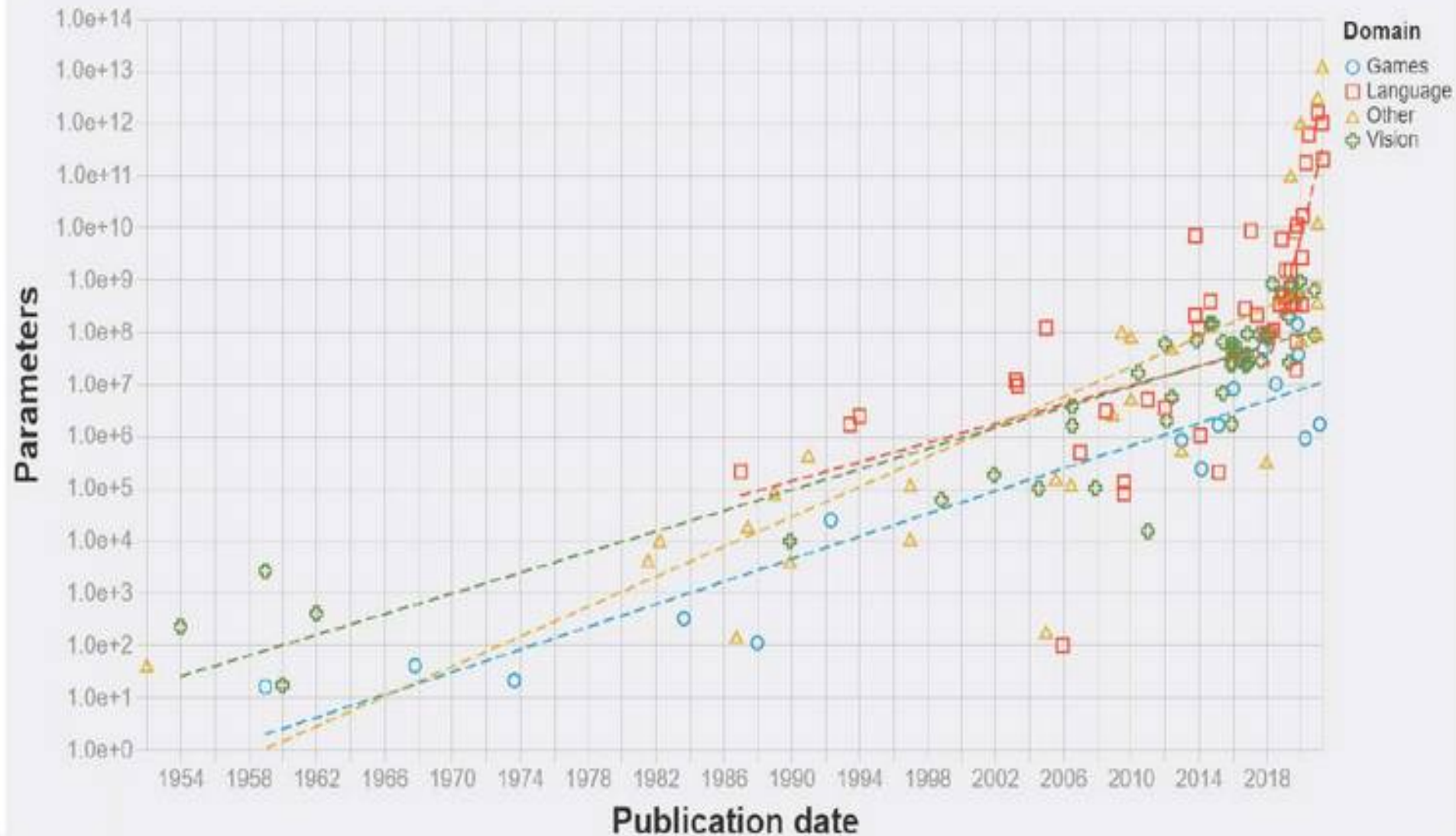
BigData and AI Algorithms

- 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,



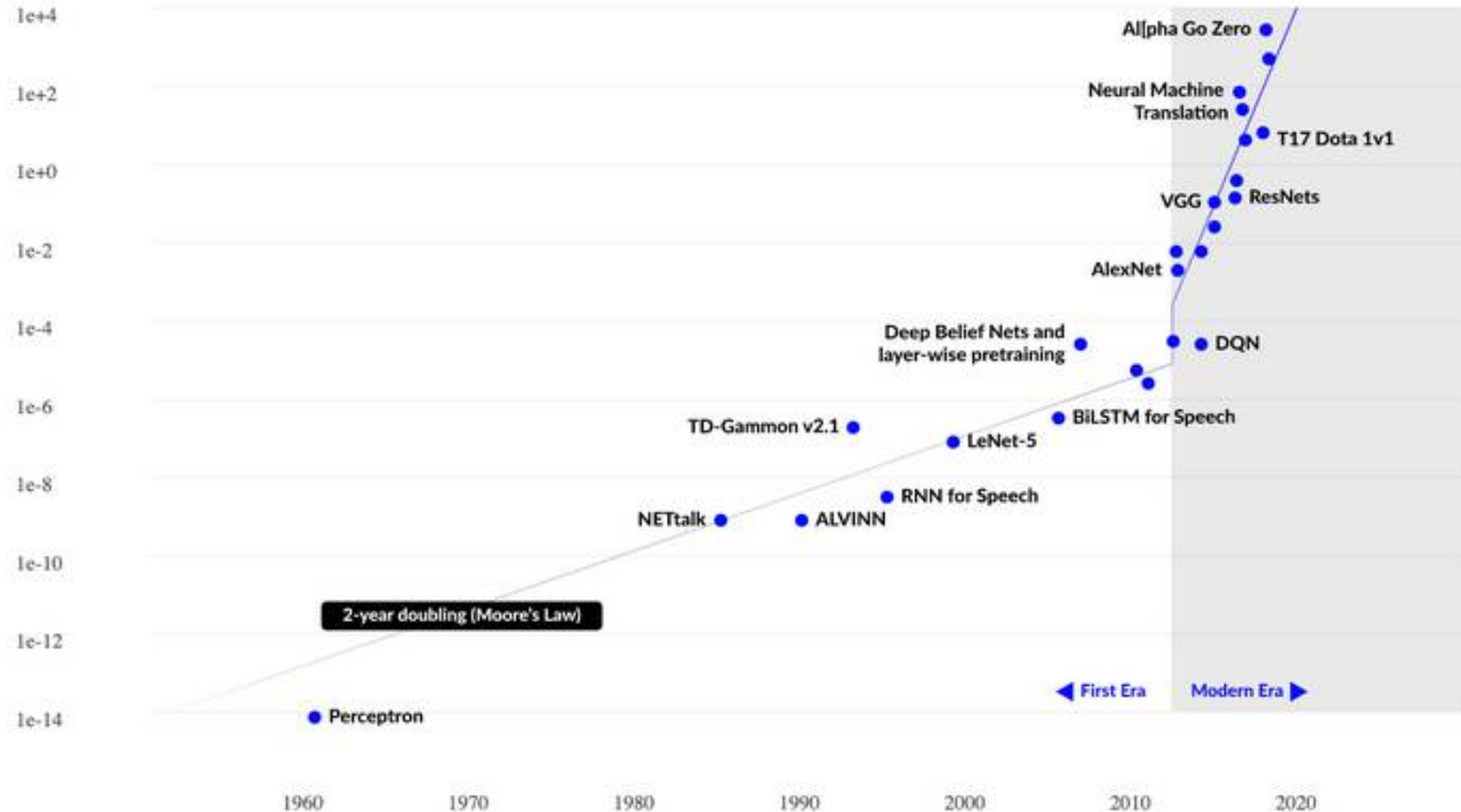
DL Relentless growth in model size

Parameter count of ML systems through time

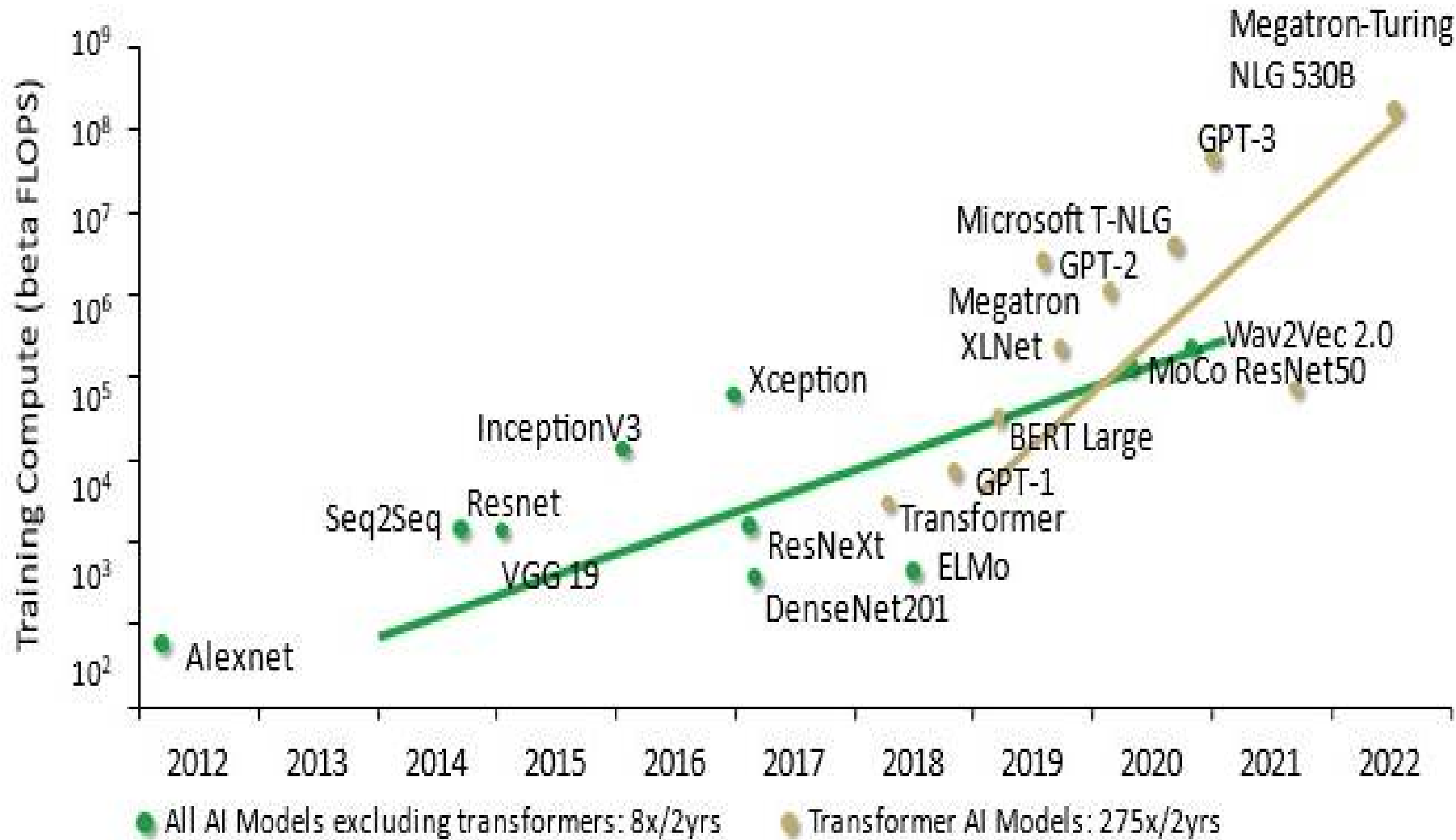


Computation Demand

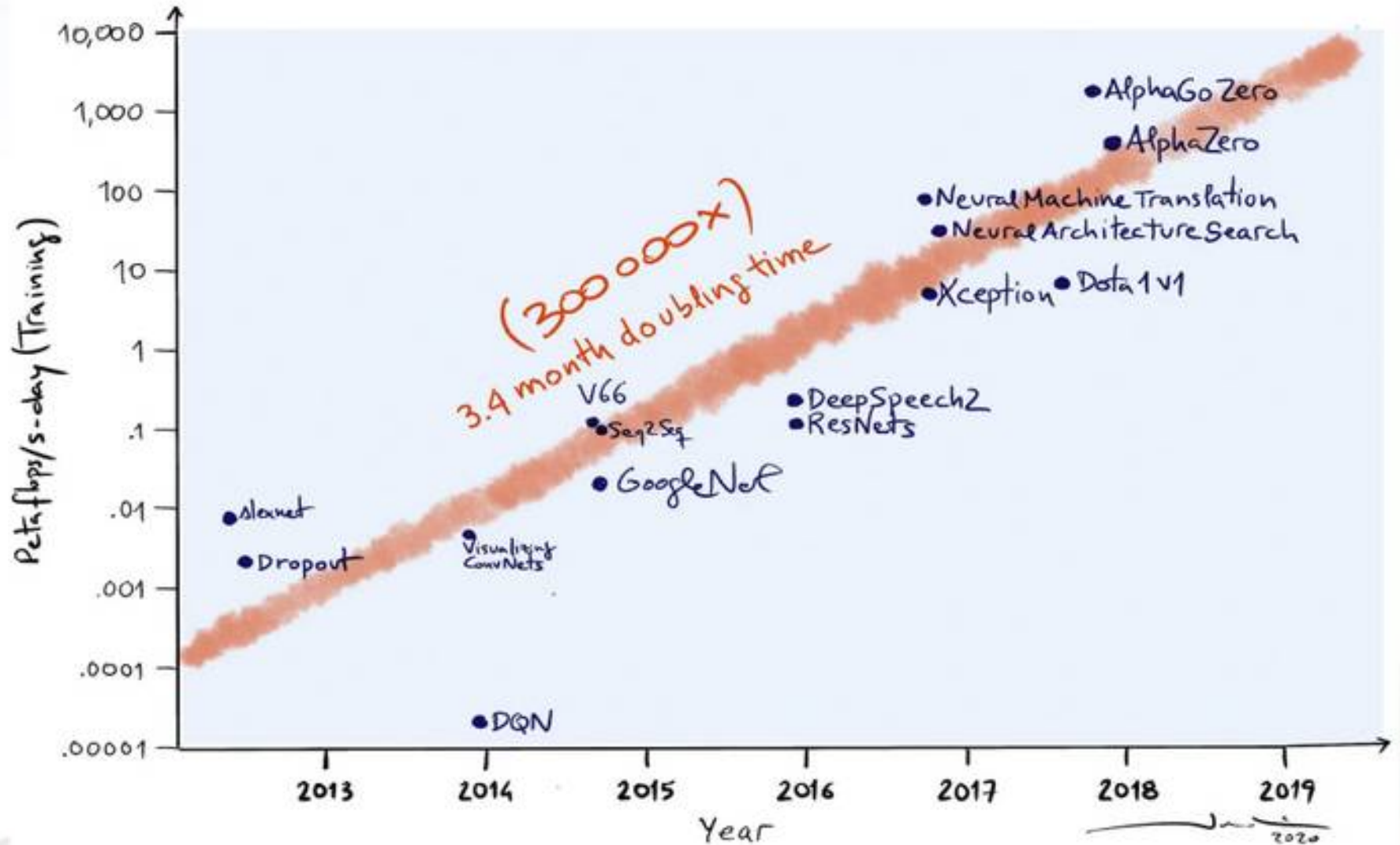
Petaflop/s-days




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



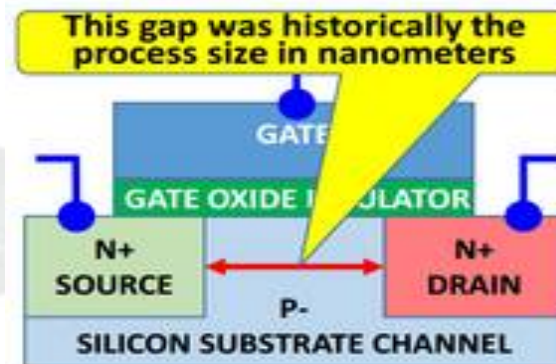
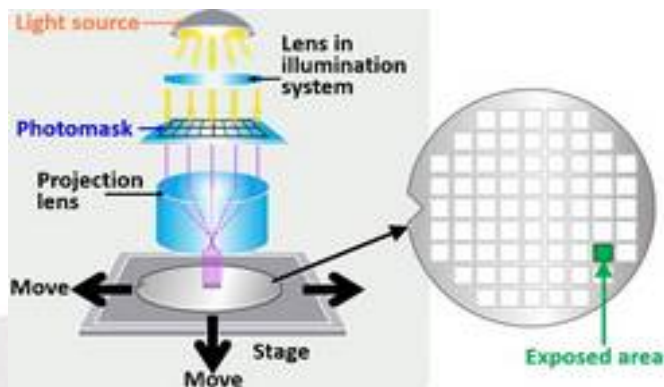
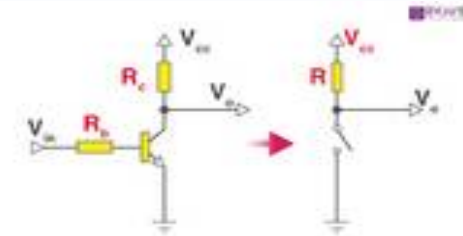
- 
- Speaker Introduction
 - Objectives of this Event
 - Mankind Progress and Industrial Revolution
 - Age of Big Data and AI
 - **Micro-electronics! Revolutionizing the World**
 - Namal Chip Design and HPC Facility

Democratization in Microelectronics

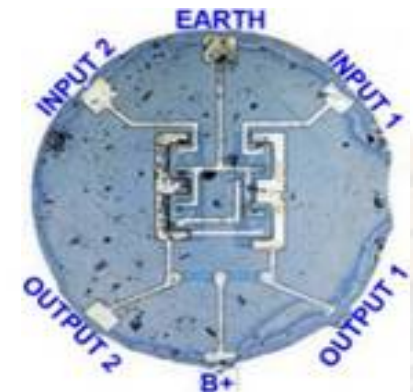
- **Microprocessors** have revolutionized automation.
- GCC has **revolutionized** the software industry.
- Linux has **revolutionized** computing industry.
- Arduino has **revolutionized** embedded computing.
- **Mathematical Models, Development Frameworks and Opensource datasets** have been **revolutionizing** the computing intellectuality.
- RISC-V is **revolutionizing** the Secure Open-source 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).



IC 1958 by Jack Kilby at Texas Instruments



Birth of Computing

Mechanical Computing by Charles Babbage 1822

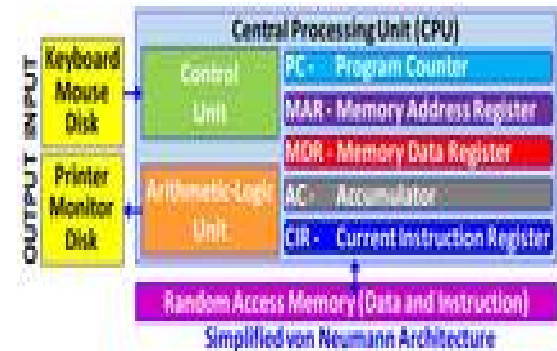
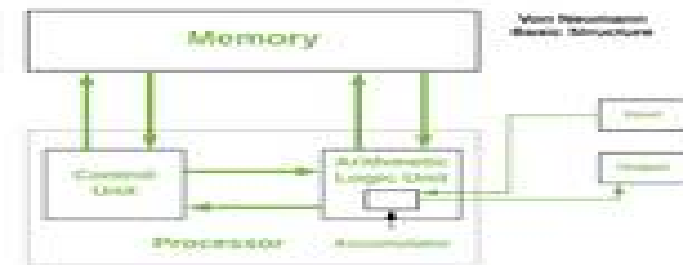
- Arch: gears, levers, and rotating shafts.
- Storage: 1K Decimal Digits
- Programming: Punch Card
- Output: Printer

Digital Computing John Von Neumann

- 1945 Electronic Discrete Variable Automatic Computer
- **1946 ENIAC (Electronic Numerical Integrator and Computer)**
- First general-purpose electronic computer by Mauchly and Eckert
- **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.

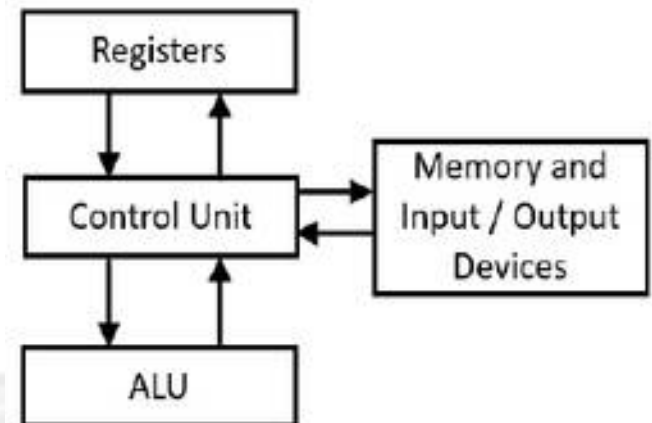
IBM PC 1981.

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

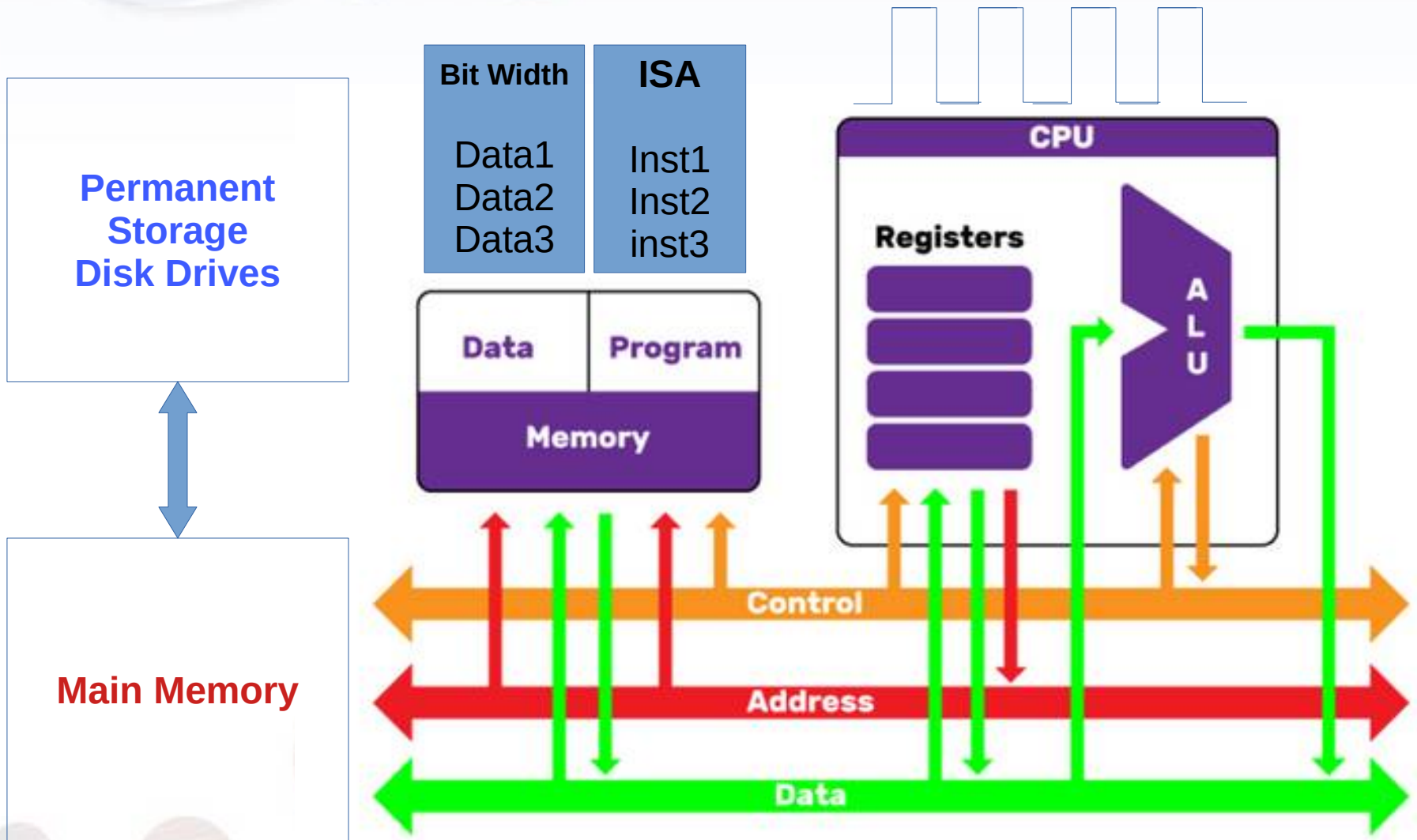


Basic Processor Architecture

- 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

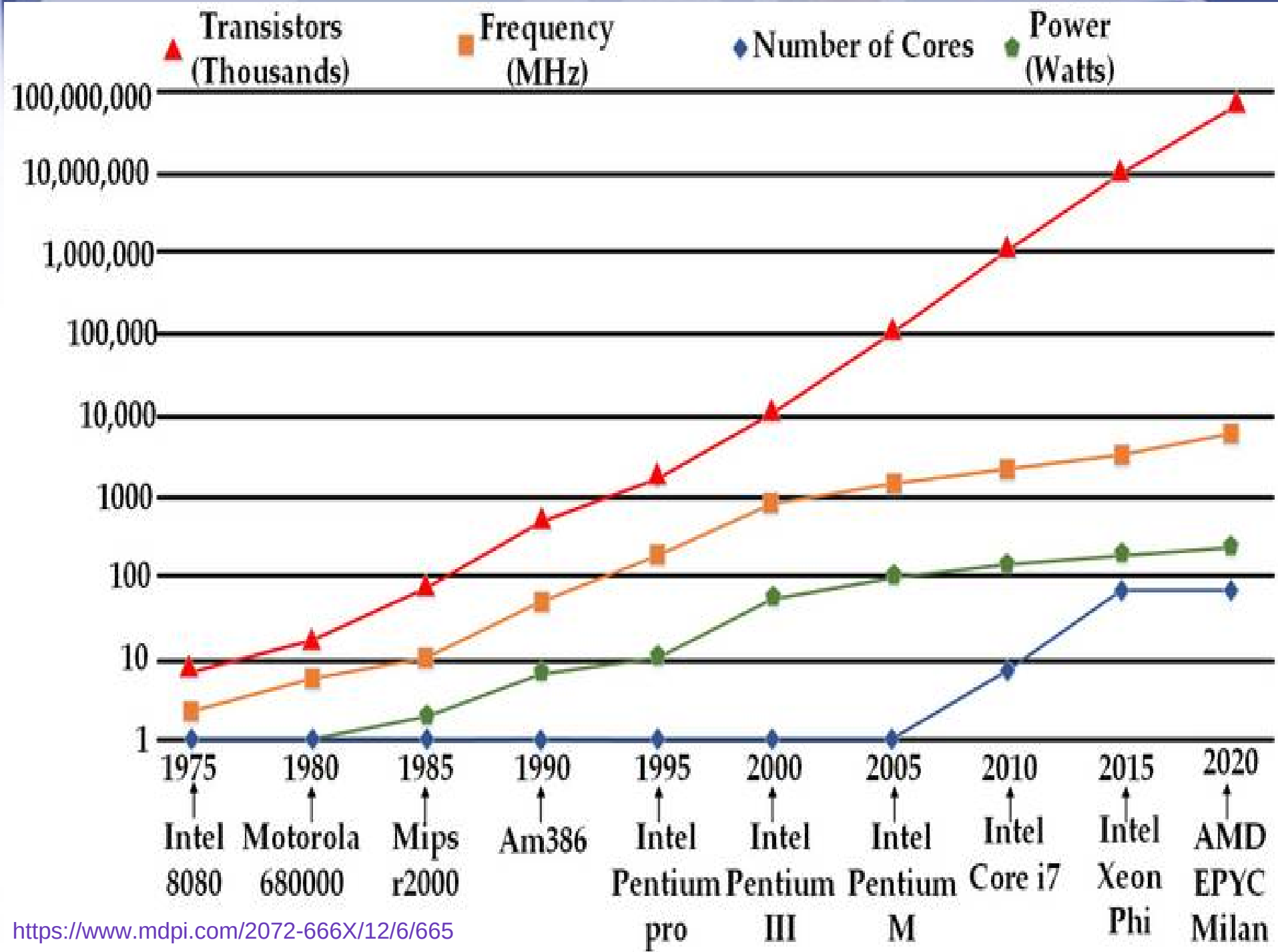
- RISCV

Vector processors

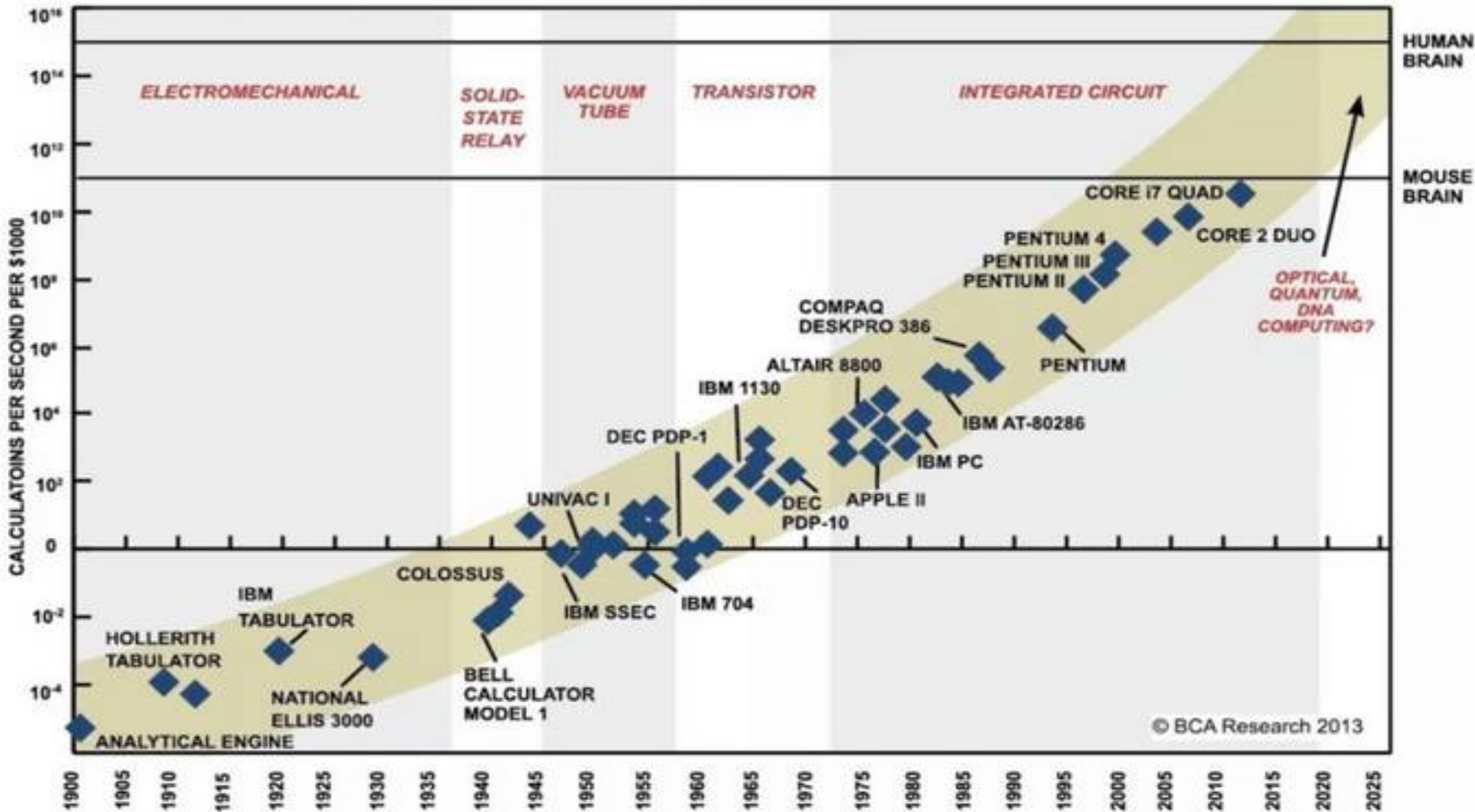
- NEC SX-6

- 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. DATAPPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

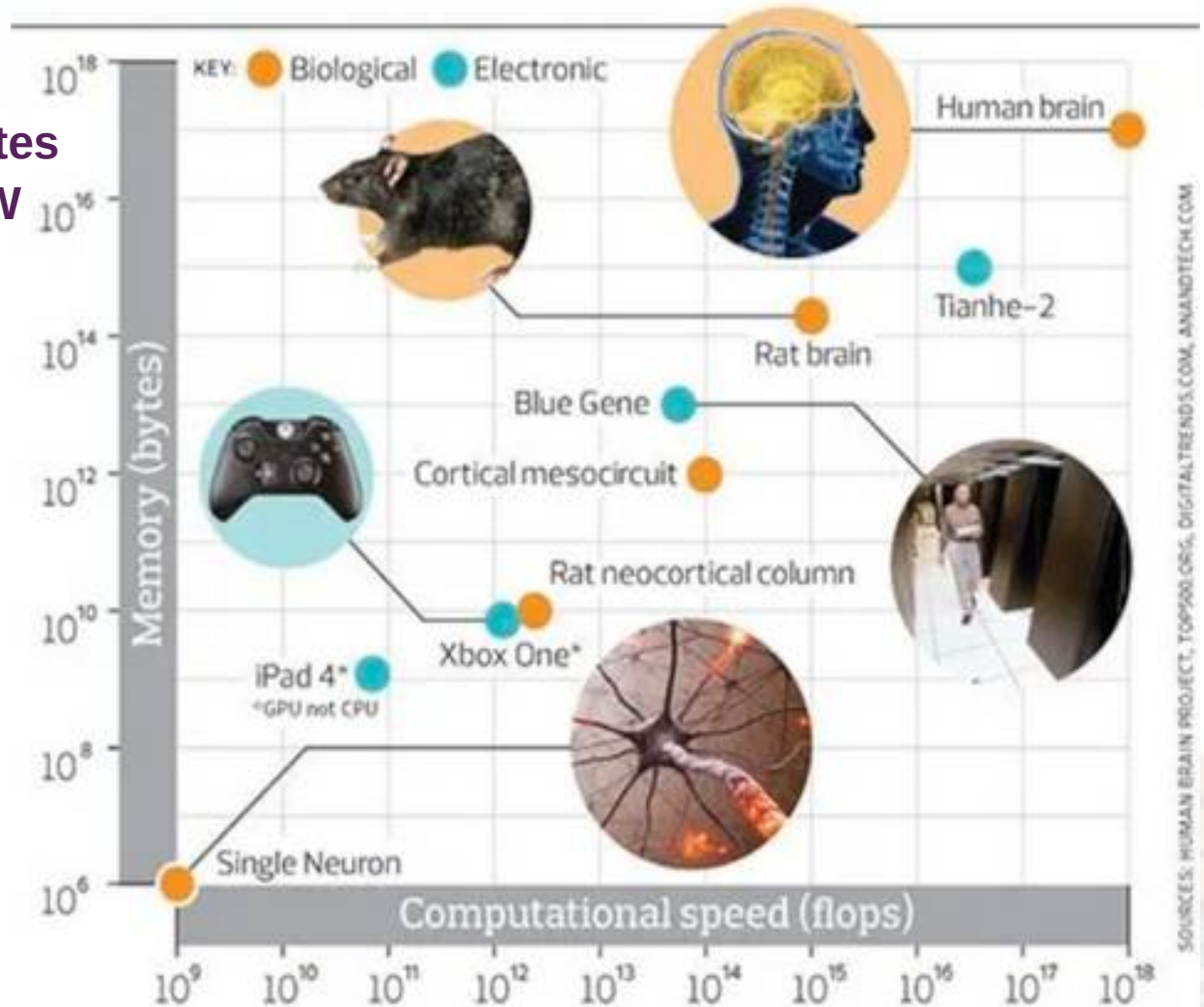
1 Operation / Second = 1 B\$

1B Operation / Second < 1\$

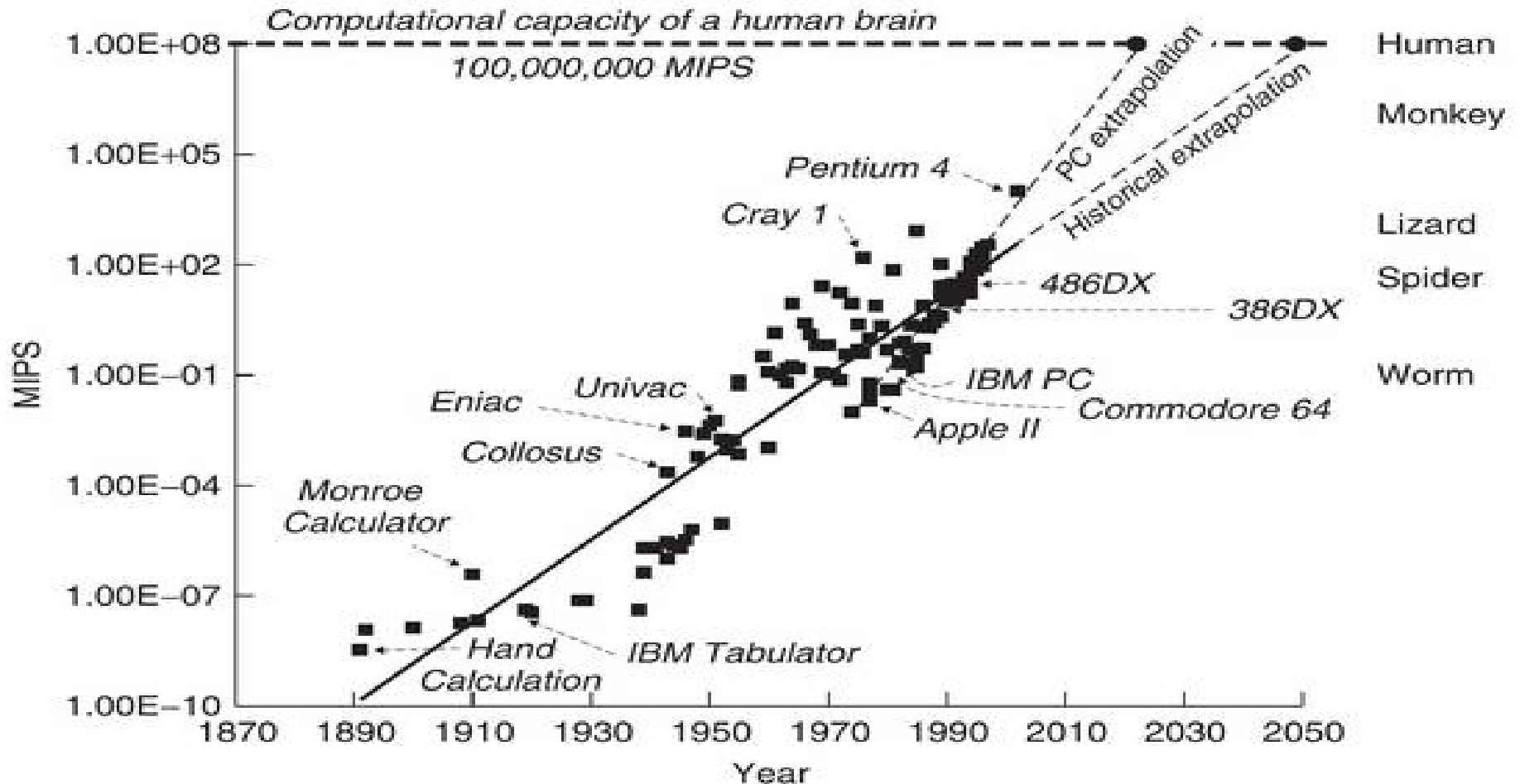
Compute Vs Intellectual Capability

● Human Brain:

- 1 Exa FLOPS
- Memory 100 Peta bytes
- 100 Tera OPS @ 10 W



Computational Capability of Processor



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

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

What kept alive the Moors Law (uProcessor)

- Technological Innovation
- Market
- Chip Industry
- Digital Data

Birth of Compilers

Ada Lovelace wrote algorithm (in 1843) 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-14.....

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:

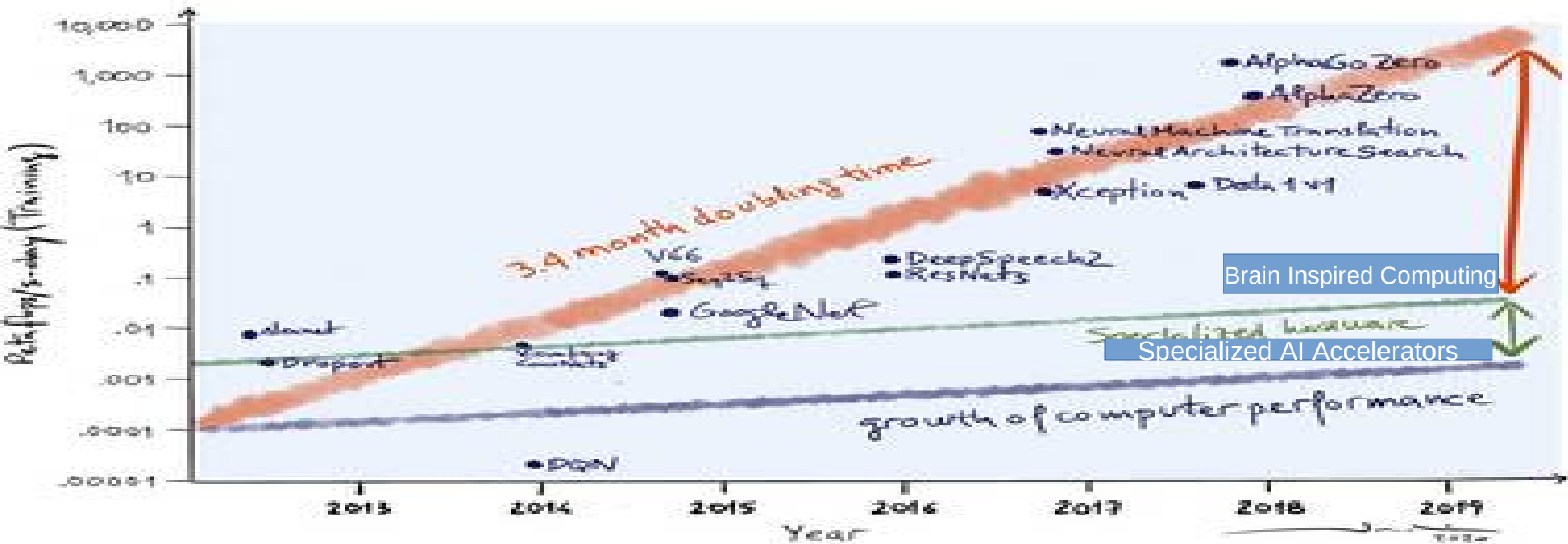
Extensions:

Community and Documentation:

Portability:

Free Software Philosophy:

AI Requirements and Processor Performance Gap

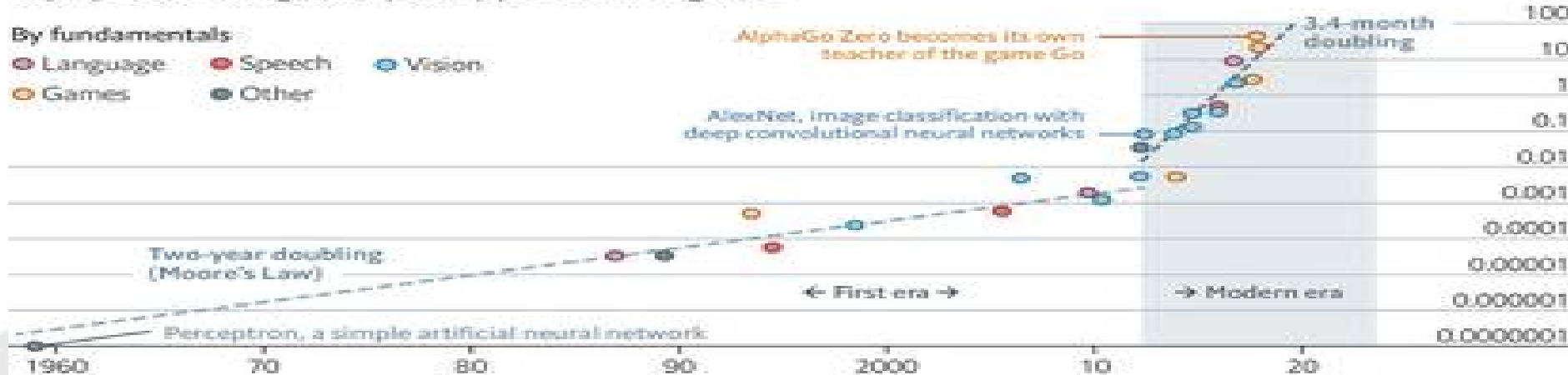


Deep and steep

Computing power used in training AI systems
 Days spent calculating at one petaflop per second*, log scale

By fundamentals:

- Language
- Speech
- Vision
- Games
- Other



Open Source: Hardware and ASIC Tools

Democratizing Hardware 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.

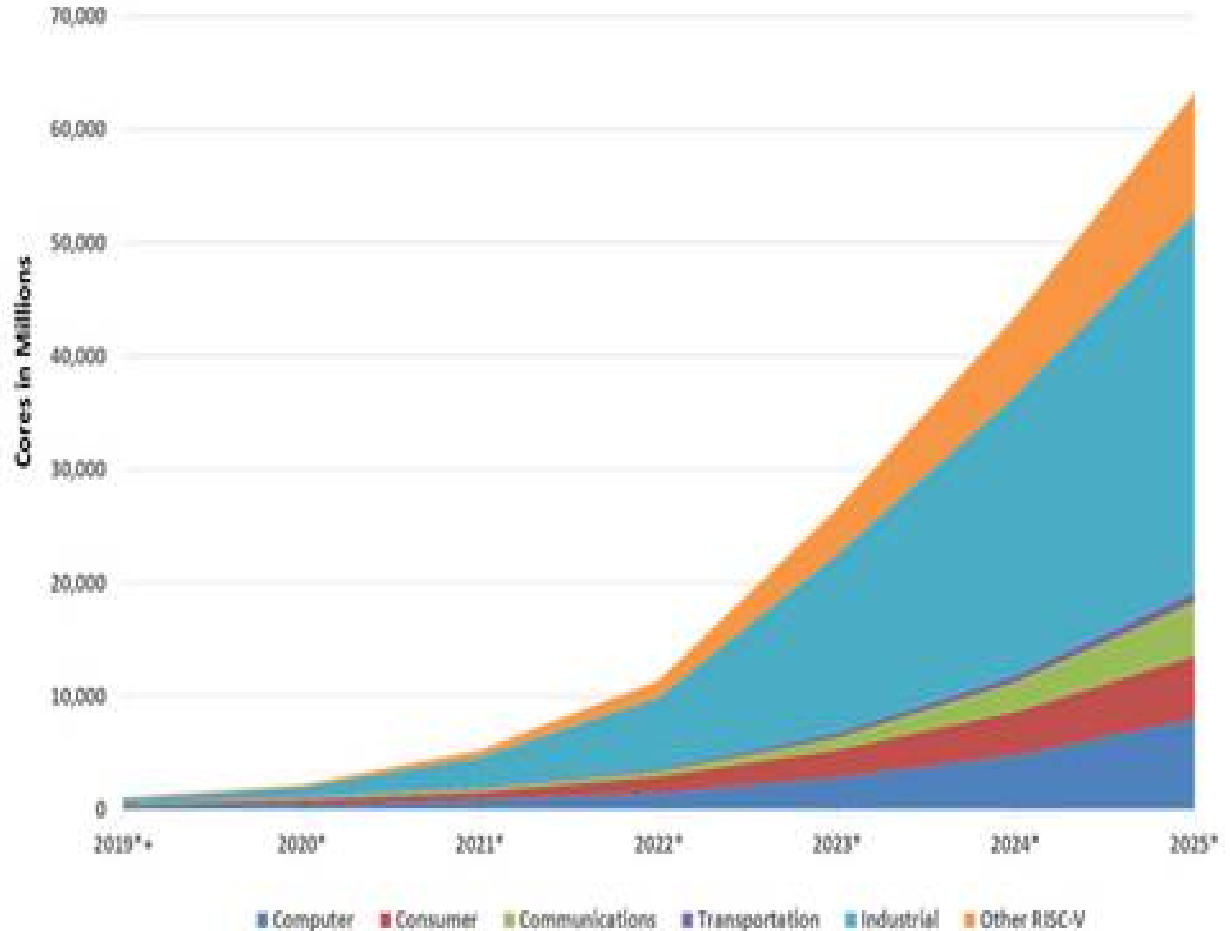
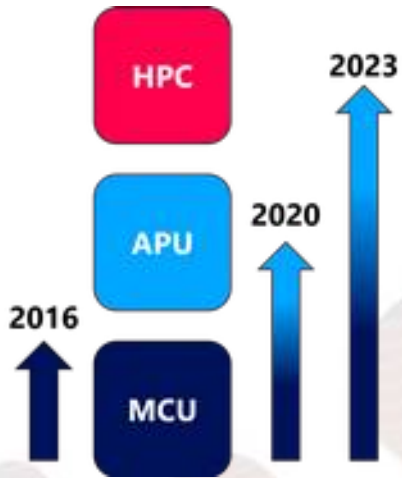
Further Customization: Expect more customized processors and domain-specific architectures tailored to niche applications.

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

Future of RISC-V



Esperanto Technologies Supercomputer-on-Chip" design, integrates many RISC-V cores on a single chip. Capable to provide 52.22 teraFLOPS



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

Source: Semico Research Corp.



Understand RISC-V Architecture

Instruction Sets


Programming Environment for RISC-V

Embedded Systems Development

Simulation and Testing

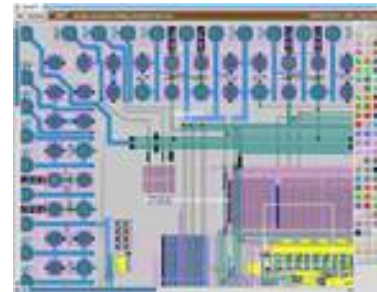
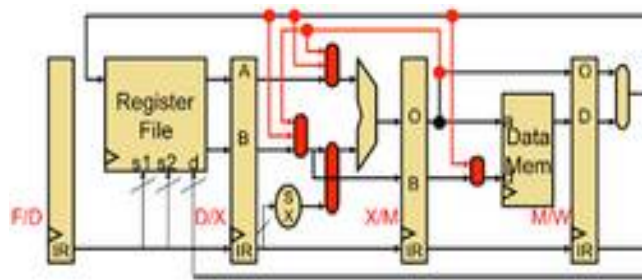
Advanced Topics in RISC-V

Industry and Research Integration

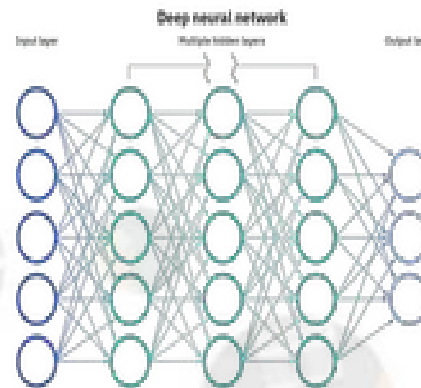
- 
- Mankind Progress
 - Age of Big Data and AI
 - Micro-electronics! Revolutionizing the World
 - **Namal Centre for AI and BigData and Electrical Engineering Department**

Centre for AI and BigData

- **OpenSource Full-Stack Ecosystem for Processor System**



- **Supercomputing for AI and BigData Applications**



OpenSource Full-Stack Ecosystem for Processor Architecture

- **Hardware Architecture**

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

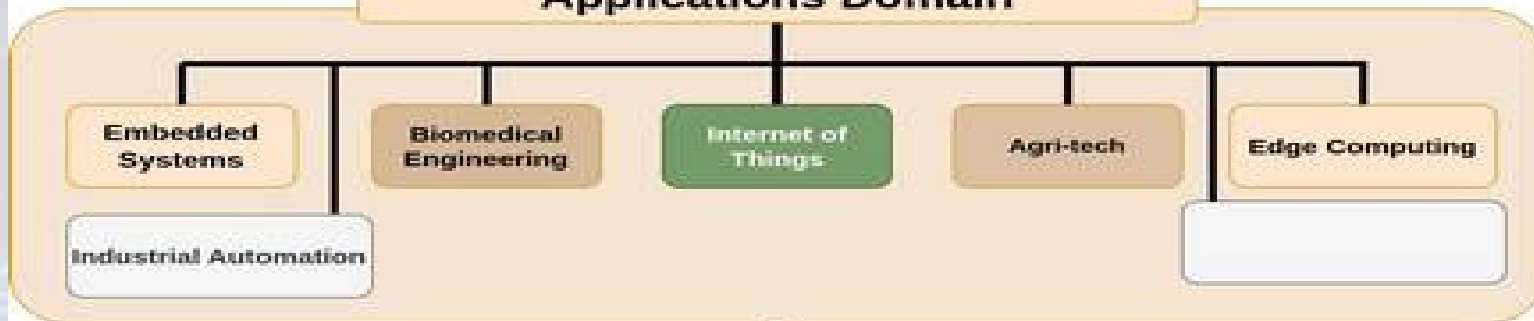
- **Single Board Computer**

- Hardware Software Co-Design
- High Performance Computing

- **Intelligent and Real-time Applications**

- Industrial Automation
- Machine Learning

Applications Domain



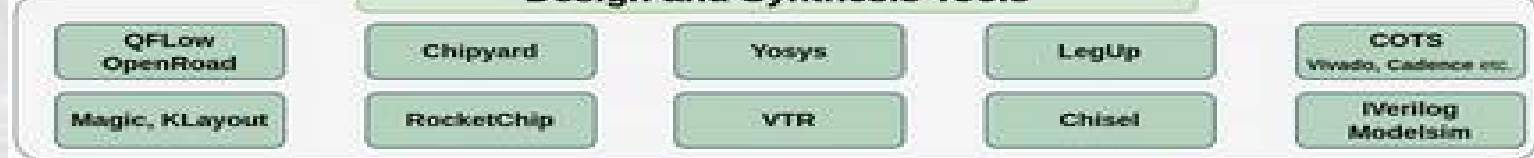
Hardware Stack



Hardware Development Approaches



Design and Synthesis Tools



Application Development Stack



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, Contrainer

Hardware System

Intelligent RACK
Infrastructure
PDU, PMS

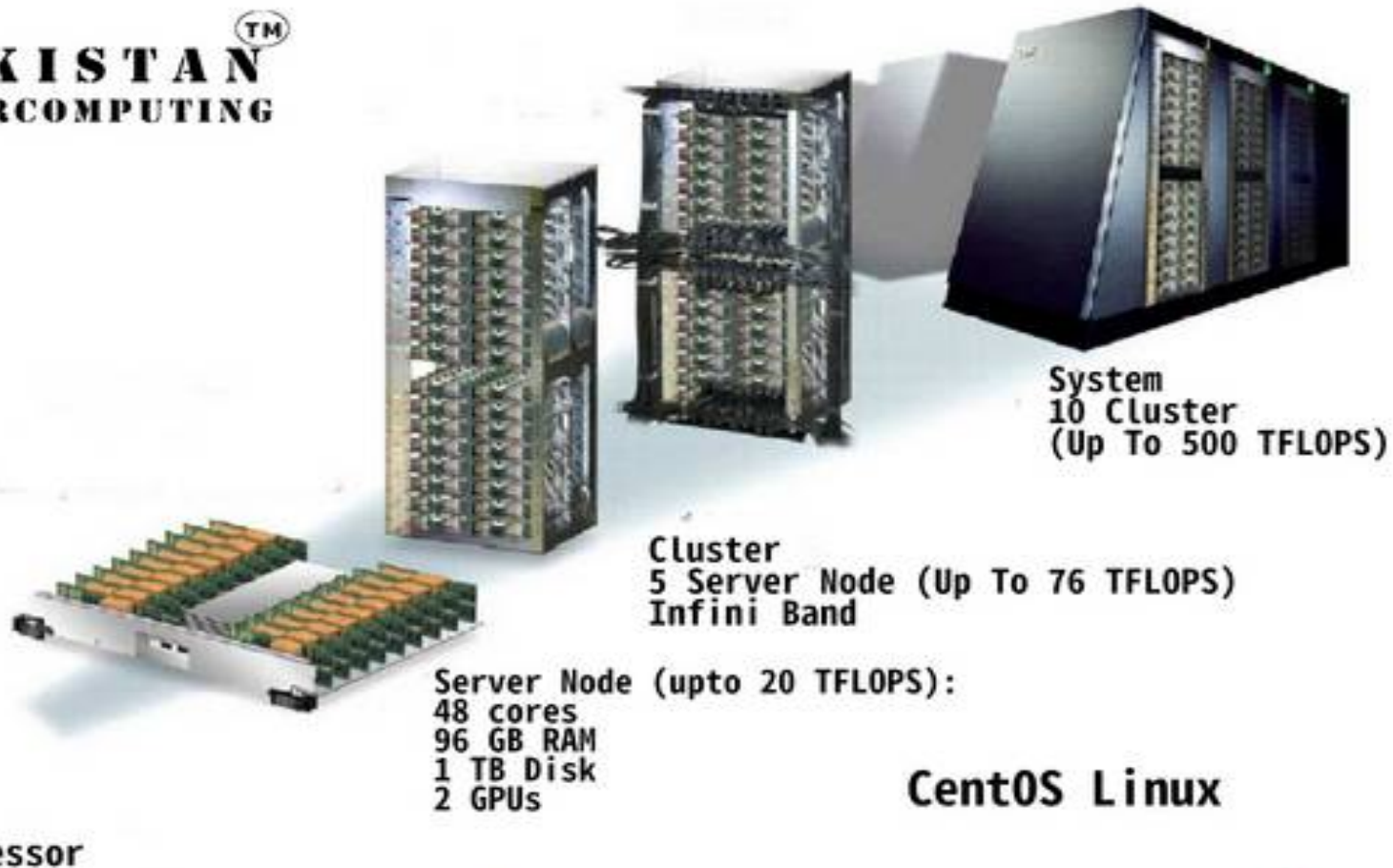
Accelerators
GPU/TPU/FPGAs

Multi-core
CISC/SuperScalar

SAN/NAS,
SSDs/NVMe

High-Speed Ethernet,
Infiniband

Developing Supercomputing for AI



**Barcelona
Supercomputing
Center**

Centro Nacional de Supercomputación



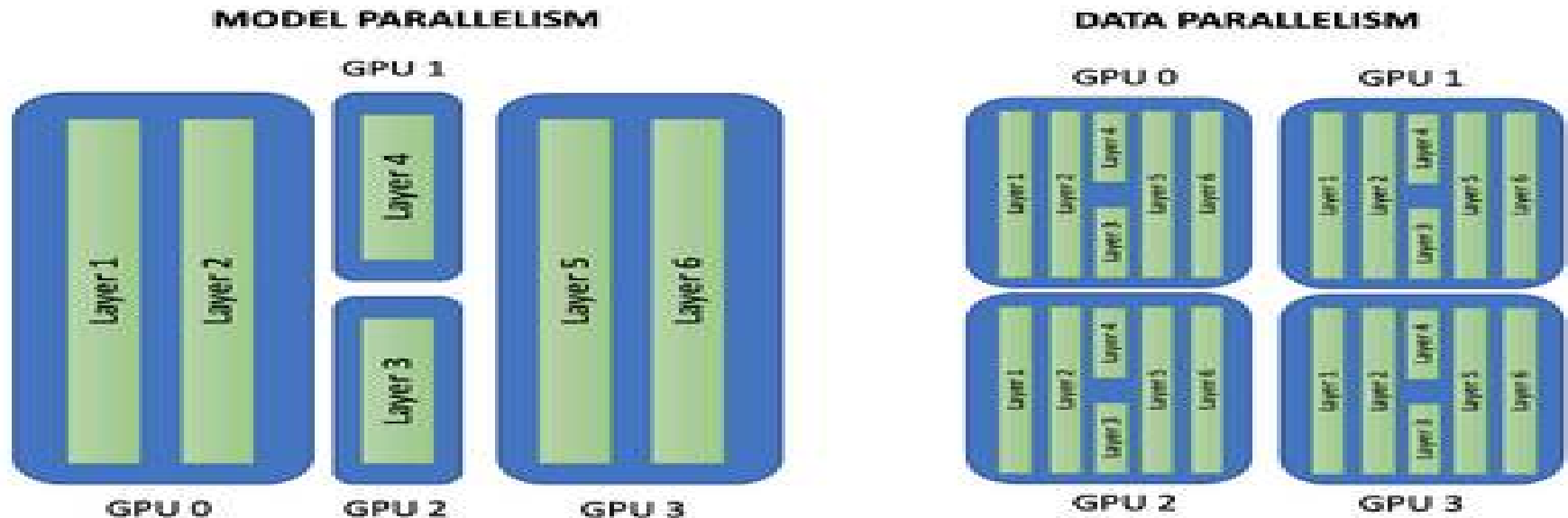
AI Model Parallelism

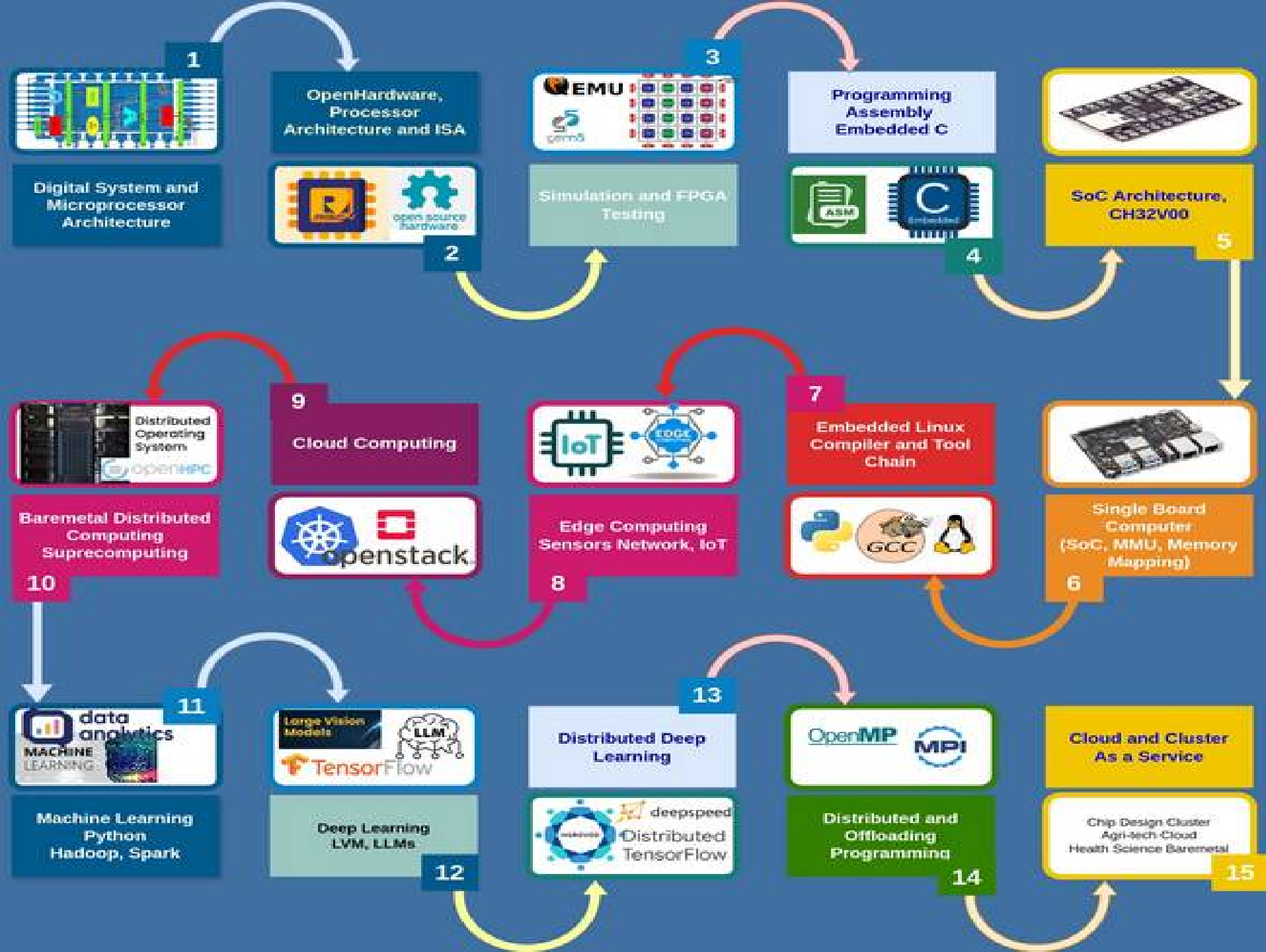
- **Model Parallelism**

Different layers of the network distributed across different devices

- **Data Parallelism**

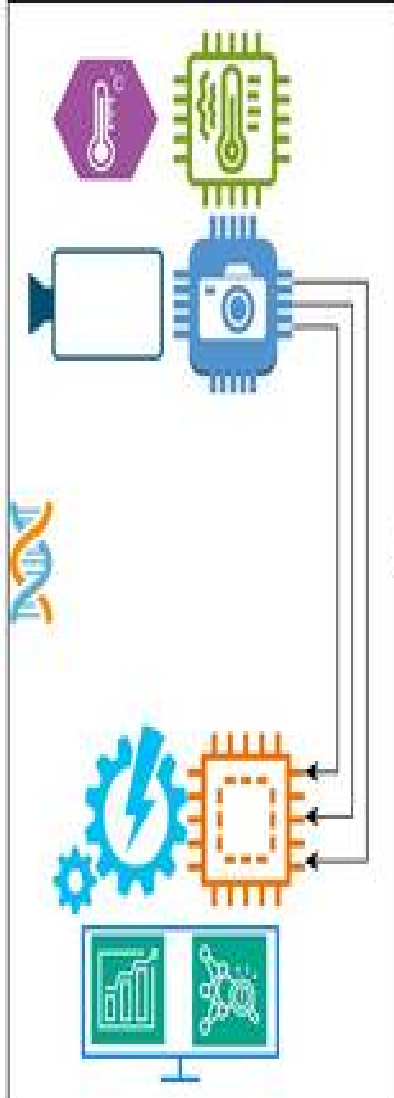
Same model in every one of the GPUs, each processing a separate piece of the data, a separate portion of the mini-batch.



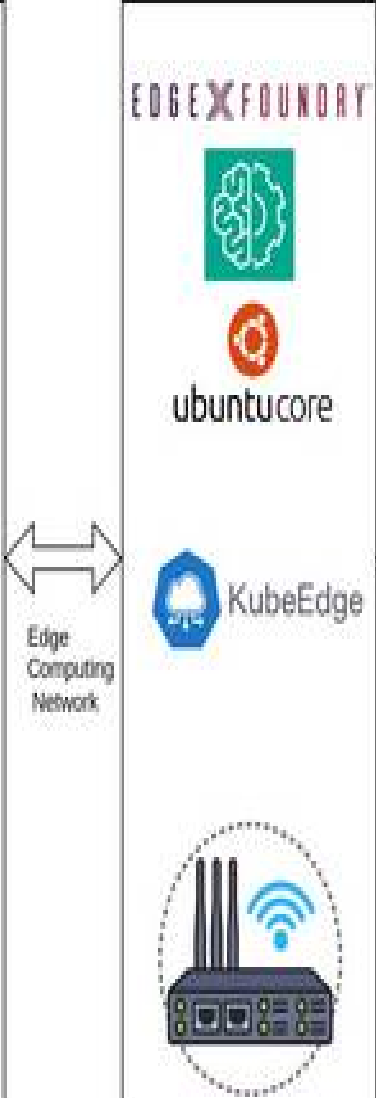




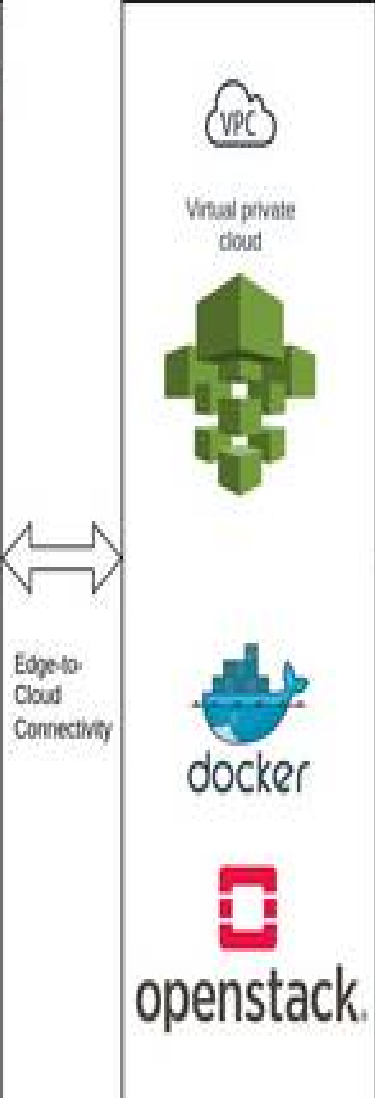
Sensors Network, IoT, Automation
Fields, Farms, Processing Units and
Research Labs



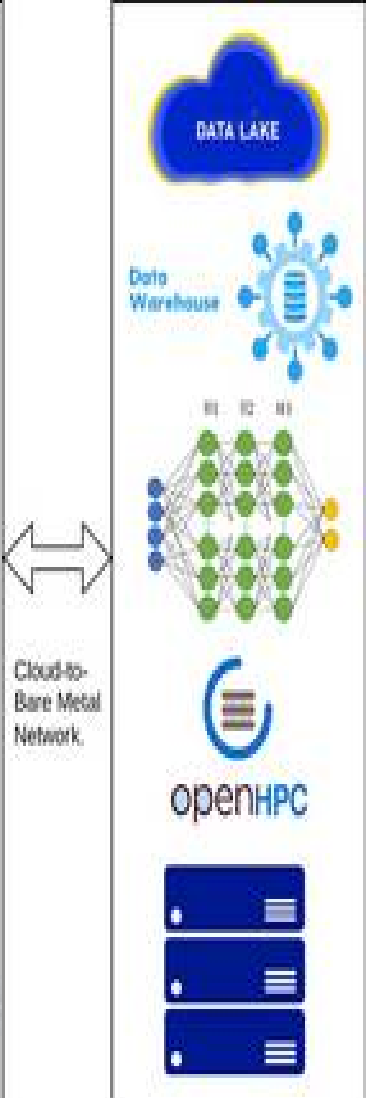
Real Time Computing
Decision Support System



Edge Computing
Basic Analysis, Pre-Processing
Weak-AI, Machine Learning



Cloud Computing
Classification, Prediction
Deep Learning



Supercomputing
HPC, Big-Data Processing
Deep/Reinforcement Learning

User Access
 Baremetal: `ssh namal-hpc@10.0.0.154`
 Cloud Application: `http://10.0.0.153:8501/`

Achievements

Bare-metal Cluster

- 1 Peta FLOPS
 - Chip Design FOSSS RTL-GDS

Cloud Applications

- Agri-Rice Classification
- Live Stock Breed Identification
- Soil Analysis
- Rehabilitation

Digital Systems

- FPGA based Computer Vision System for Rice Color Sorting
- Foot Analytics

Collaboration Areas

- a) Specialized Training Programs
- b) Research and Publications**
- c) Technology Development and Transfer

Collaboration Support

- **Technical Project**

- Local Problem having end-user

- **Master Program**

- Student Support

- Partially Support

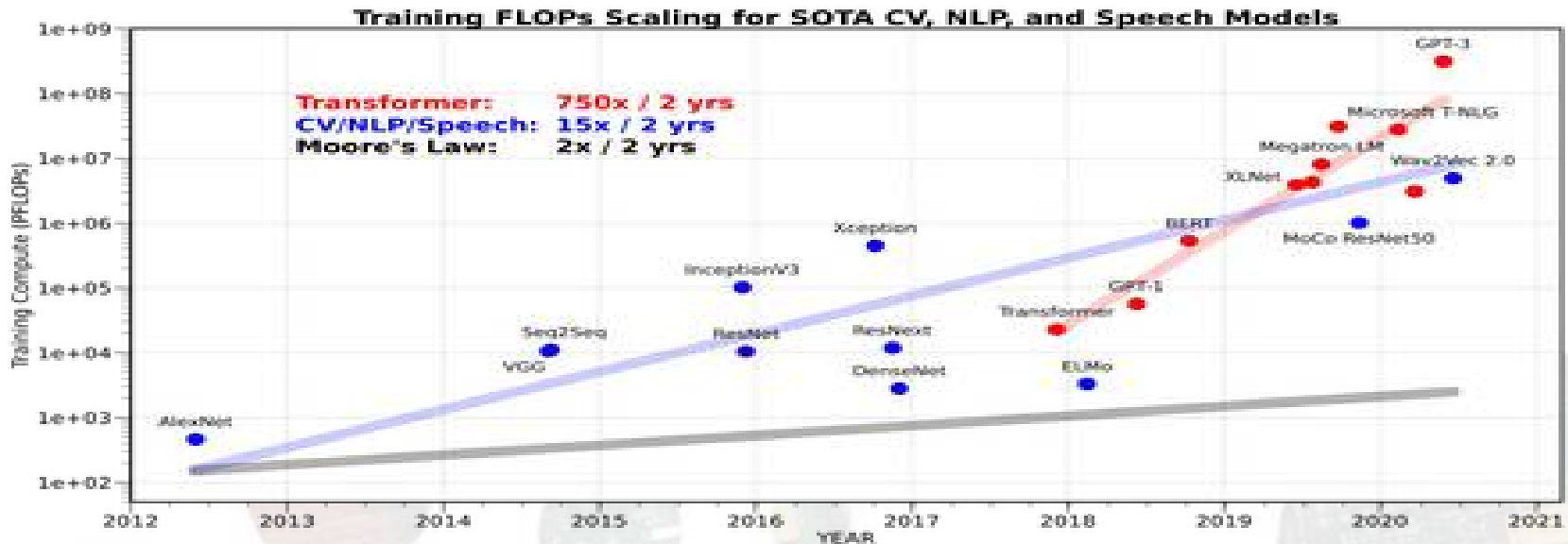
- Research Thesis in Collaboration with Industry

- Full Support

- Work under supervision of industrial and university Supervisor

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.

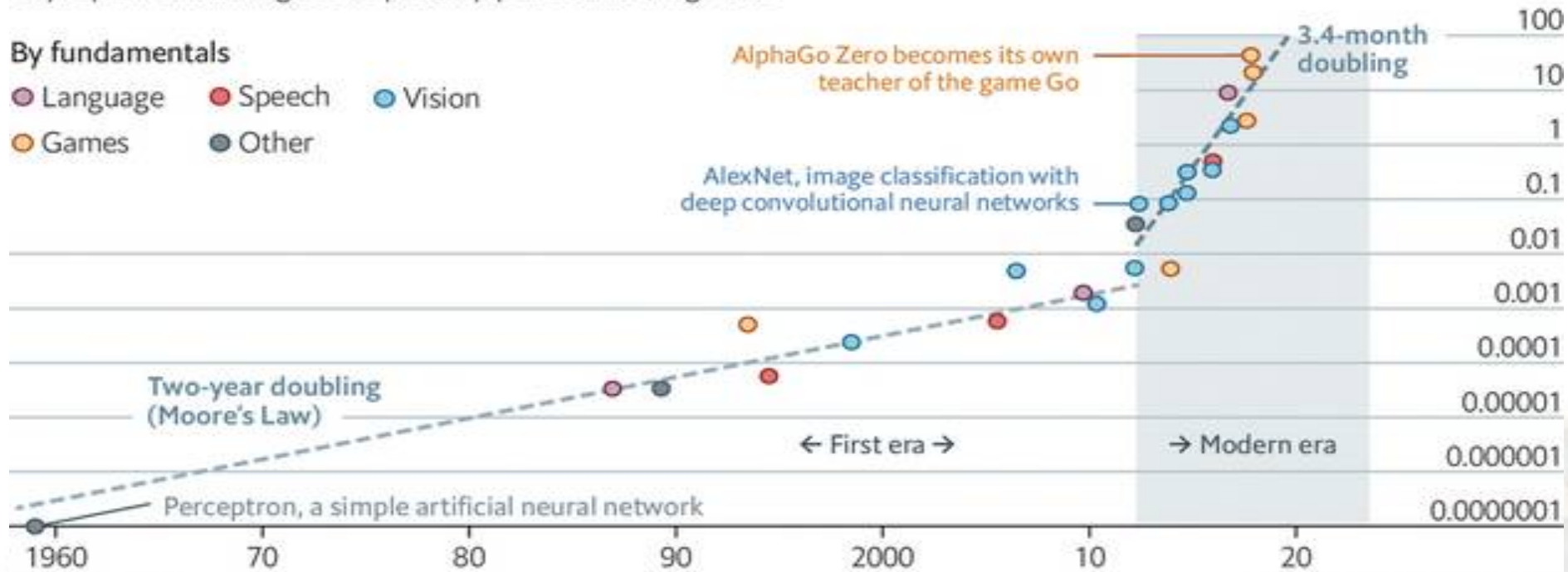
Deep and steep

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

Microprocessor based Systems: Revolutionizing the World

by: **Tassadaq Hussain**

Director Centre for AI and BigData

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