INDIAN INSTITUTE OF INFORMATION TECHNOLOGY

Dharwad

ज्ञानेन विकासः

Systems and Data Science

Dr. Animesh Chaturvedi

Assistant Professor: IIIT Dharwad

Post Doctorate: King's College London & The Alan Turing Institute

PhD: IIT Indore MTech: IIITDM Jabalpur

Indian Institute of Technology Indore

भारतीय प्रौद्योगिकी संस्थान इंदौर

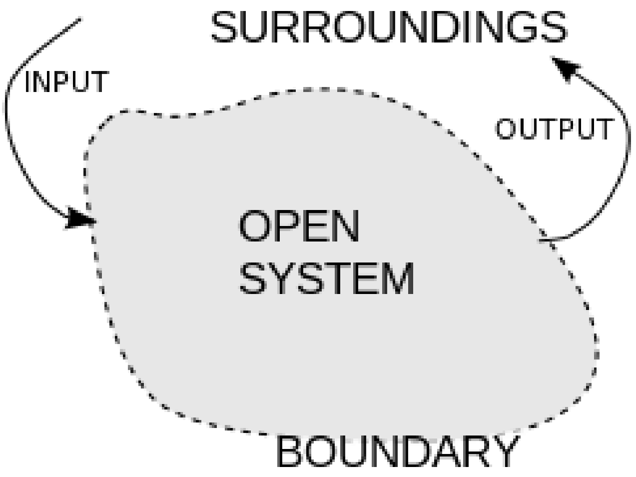
KING'S College LONDON

PDPM Indian Institute of Information Technology, Design and Manufacturing, Jabalpur

The Alan Turing Institute

Systems

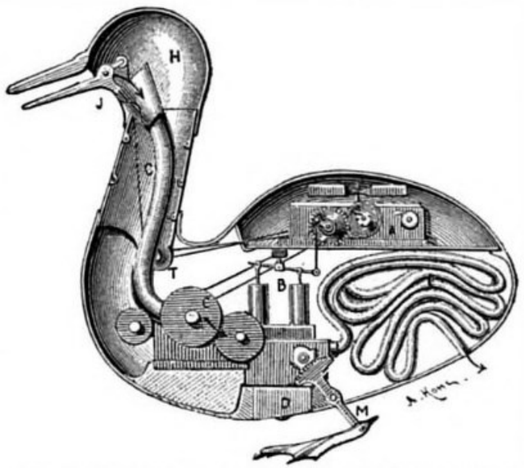
- An interdisciplinary domain
- Composed of many components (or entities) interacting with each other.
  - power grid, transportation or communication systems,
  - social and economic organizations (like cities),
  - organisms, a living cell, the human brain, and
  - an ecosystem, climate, entire universe.
- Behavior is hard to model with dependencies, relationships,
  - interactions between their components or
  - interactions between system and its environment



https://en.wikipedia.org/wiki/Complex_system

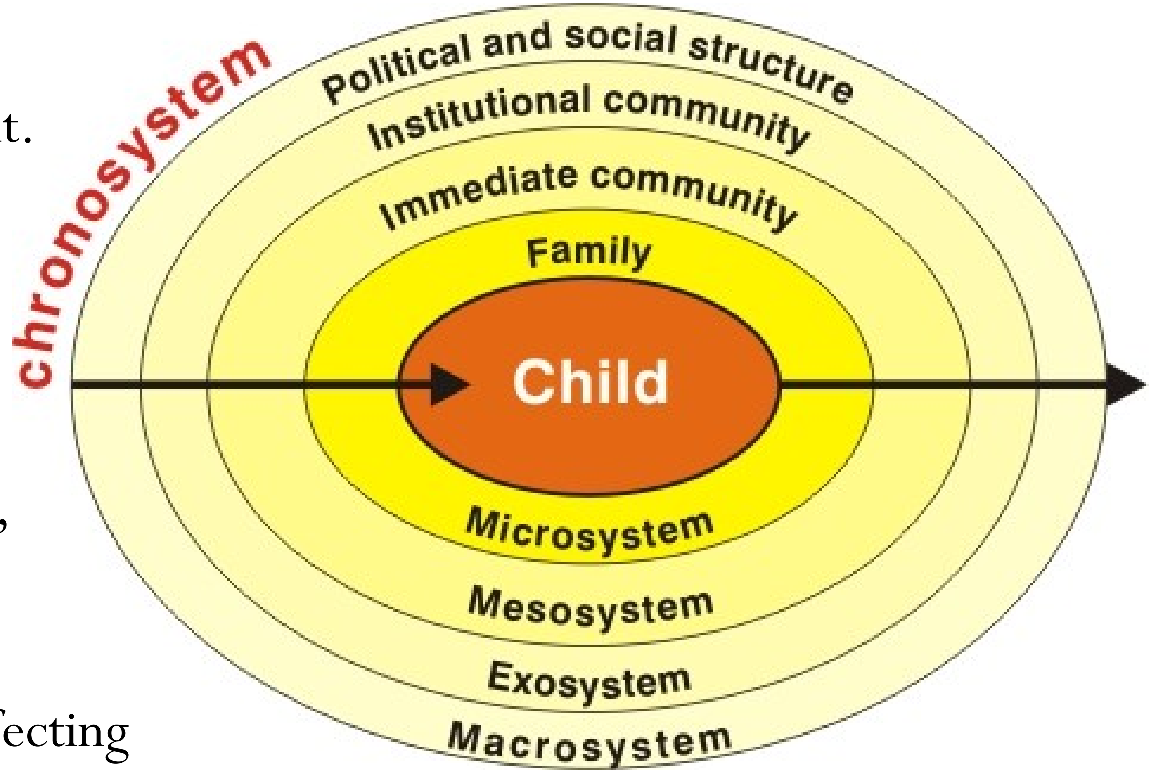
Systems and Reductionism

- Reductionism is old domain since 16th century
- explains system in terms of parts and their interactions
  - Define a domain of possible parts
  - Generate inputs over the interaction between parts
  - Perform a deterministic computation on the input data
  - Aggregate the results
- interprets a complex system as the sum of its parts



Systems theory and Data

- Micro-system:
  - The system closest to the client.
- Meso-system:
  - Relationships among systems in an environment.
- Exo-system:
  - A relationship between two systems that has an indirect effect on a third system.
- Macro-system:
  - A larger system that influences clients, policies, administration, and culture.
- Chrono-system:
  - A system composed of significant life events affecting adaptation.



https://en.wikipedia.org/wiki/Systems_theory

Systems theory and Data

- Interdisciplinary study of interconnected component or entities in **System data**
- Component (or entities) can be natural or machine.
- System has following properties
  - bounded by space and time.
  - influence by its environment.
  - structure, purpose, and functioning.
- System expresses **Synergy** (means working together).
- System expresses **Emergence** (means properties or behaviors appears only when the component or entities interact).

https://en.wikipedia.org/wiki/Systems_theory

Systems theory and Data

- Systems theory model a system's dynamics, constraints, conditions, principles (purpose, measure, methods, tools), and resource optimization.
- Change in a system's component or entities may affect other entities or the whole system.
- **Predict changes** in patterns of behavior based on **data analysis**.
- **System learns and adapts** with its environment based on **system data**.
- Systems support and maintain other systems to prevent failure.
- Dynamic or Active systems has components that interact in behaviors and processes.
- Passive systems has components that are being processed.
- Example
  - a program/code file is passive and
  - same code is a active process executing on RAM and CPU.

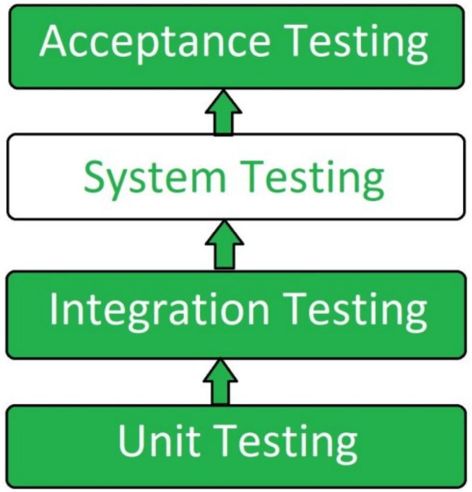
https://en.wikipedia.org/wiki/Systems_theory

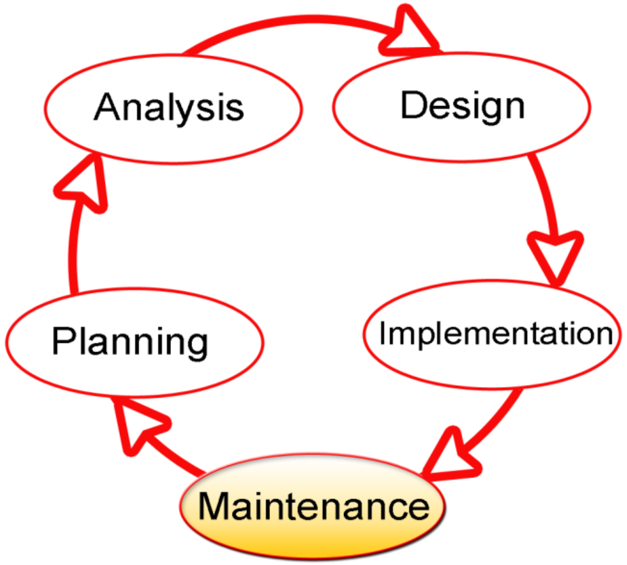
System properties: Data patterns

- **Nonlinearity:** a set of simultaneous equations with the variables of a polynomial degree higher than one
- **Emergence:** entities have properties emerging only when the they interact together
- **Spontaneous order** arises from interactions between component of disordered system
- **Adaptation:** behavior of mutation and self-organization for the changes
- **Feedback loops:** outputs are routed back as inputs as part of a chain of cause-and-effect that forms a circuit or loop

Systems Engineering Definitions

- "…a robust approach to the design, creation, and operation of systems. … identification and quantification of system goals, creation of alternative system design concepts, performance of design trades, selection and implementation of the best design, verification that the design is properly built and integrated, and post-implementation assessment …" *NASA Systems Engineering Handbook, 1995.* [3]

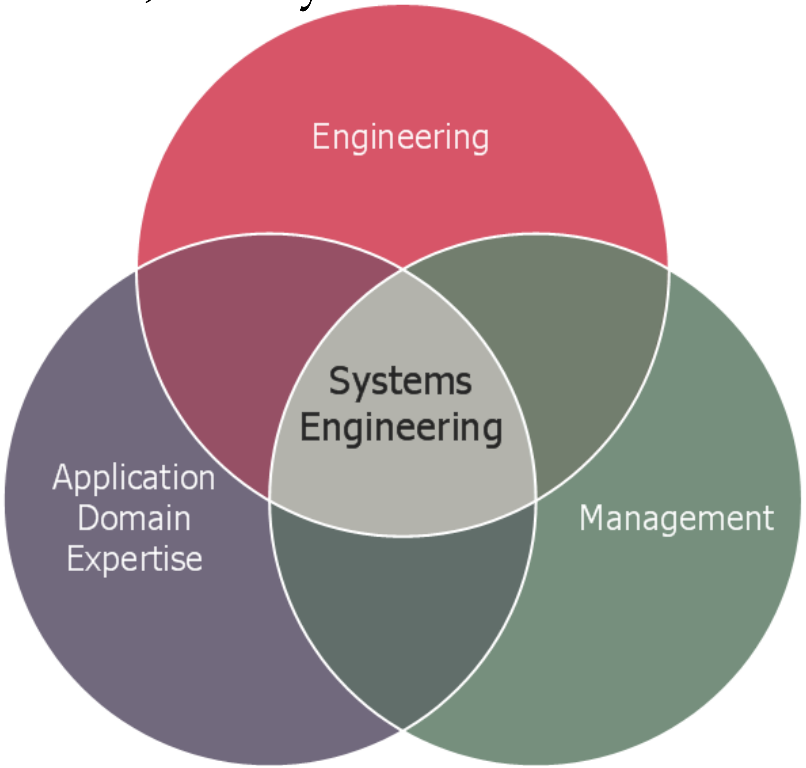




[3] *NASA Systems Engineering Handbook*. NASA. 1995. SP-610S.
https://en.wikipedia.org/wiki/Systems_engineering

Systems Engineering Definitions

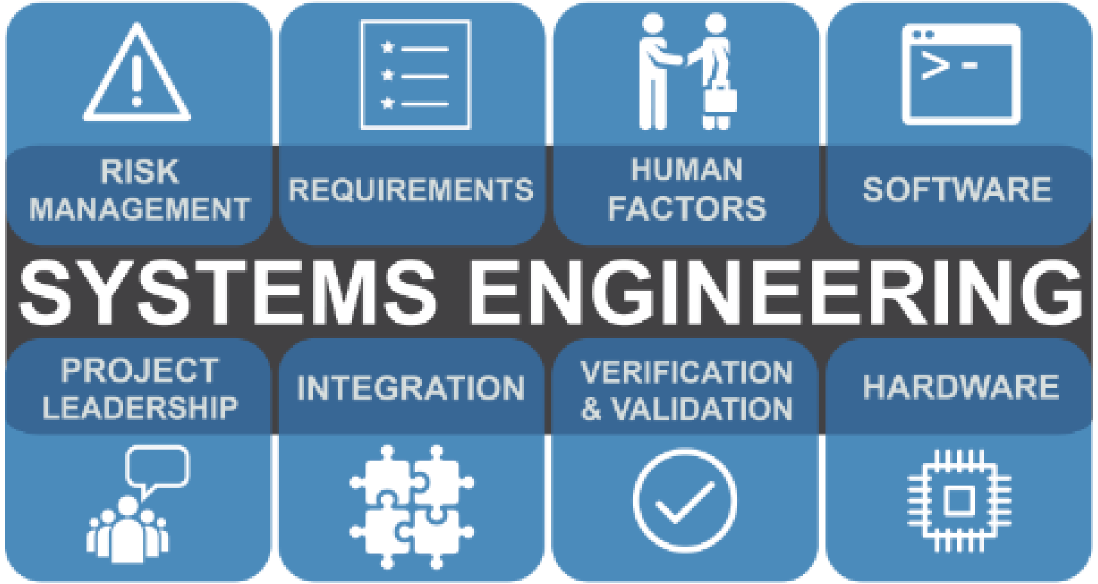
- "…a system is an integrated composite of people, products, and processes that provide a capability to satisfy a stated need or objective." [4]
- "Systems engineering consists of two significant disciplines: the technical knowledge domain in which the systems engineer operates, and systems engineering management." [4]



[4] *Systems Engineering Fundamentals*. Defense Acquisition University Press, 2001

Systems Engineering Definitions

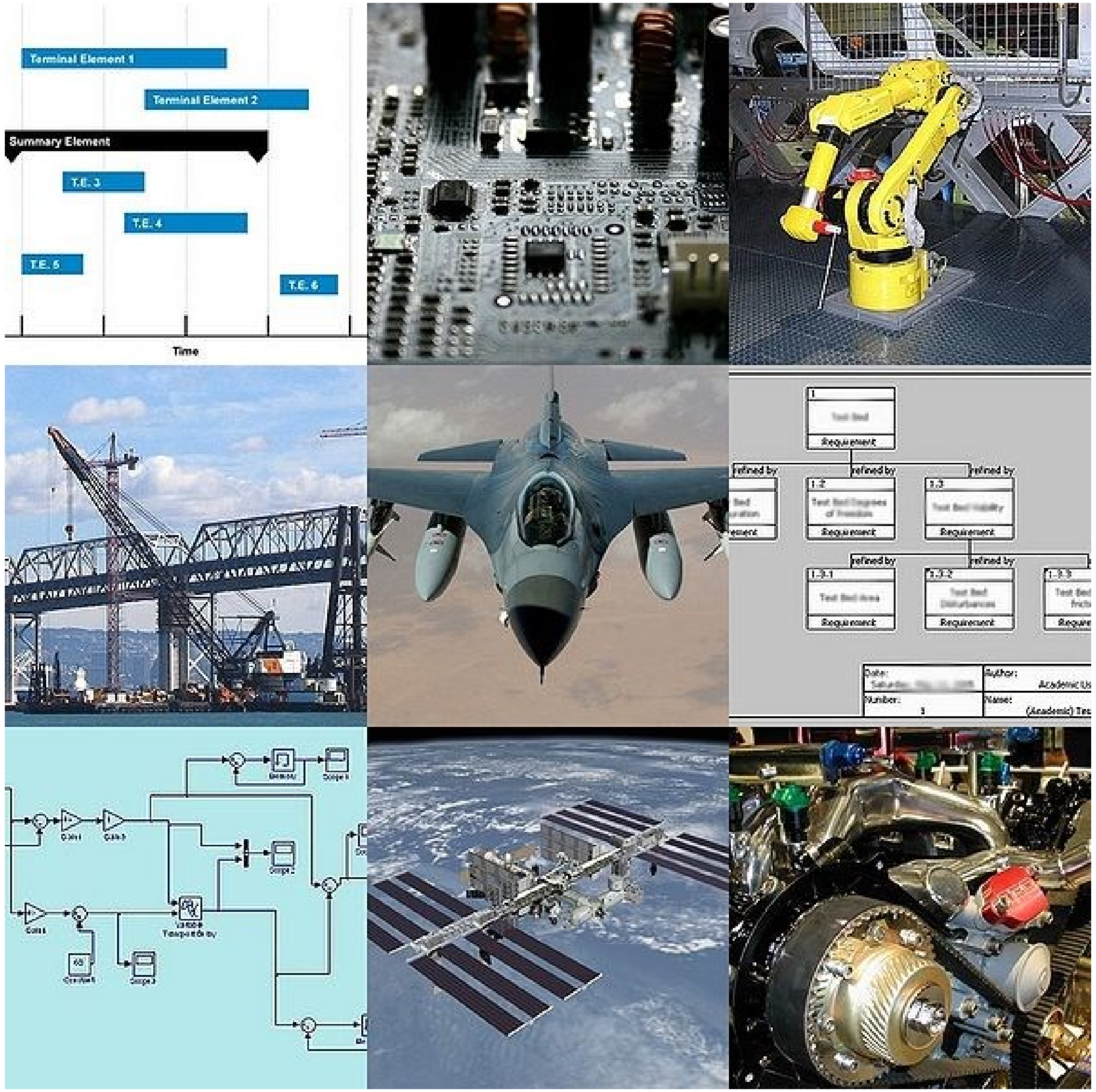
- "An interdisciplinary approach and means to enable the realization of successful systems" *INCOSE handbook, 2004.* [5]
- "…a branch of engineering which concentrates on the design and application of the whole as distinct from the parts, looking at a problem in its entirety, taking account of all the facets and all the variables and linking the social to the technological." *Conquering Complexity, 2005.* [6]



[5] *Systems Engineering Handbook*, version 2a. INCOSE. 2004.
[6] *Conquering Complexity: lessons in defence systems acquisition*, The Defence Engineering Group. University College London. 2005.

Systems Engineering Examples

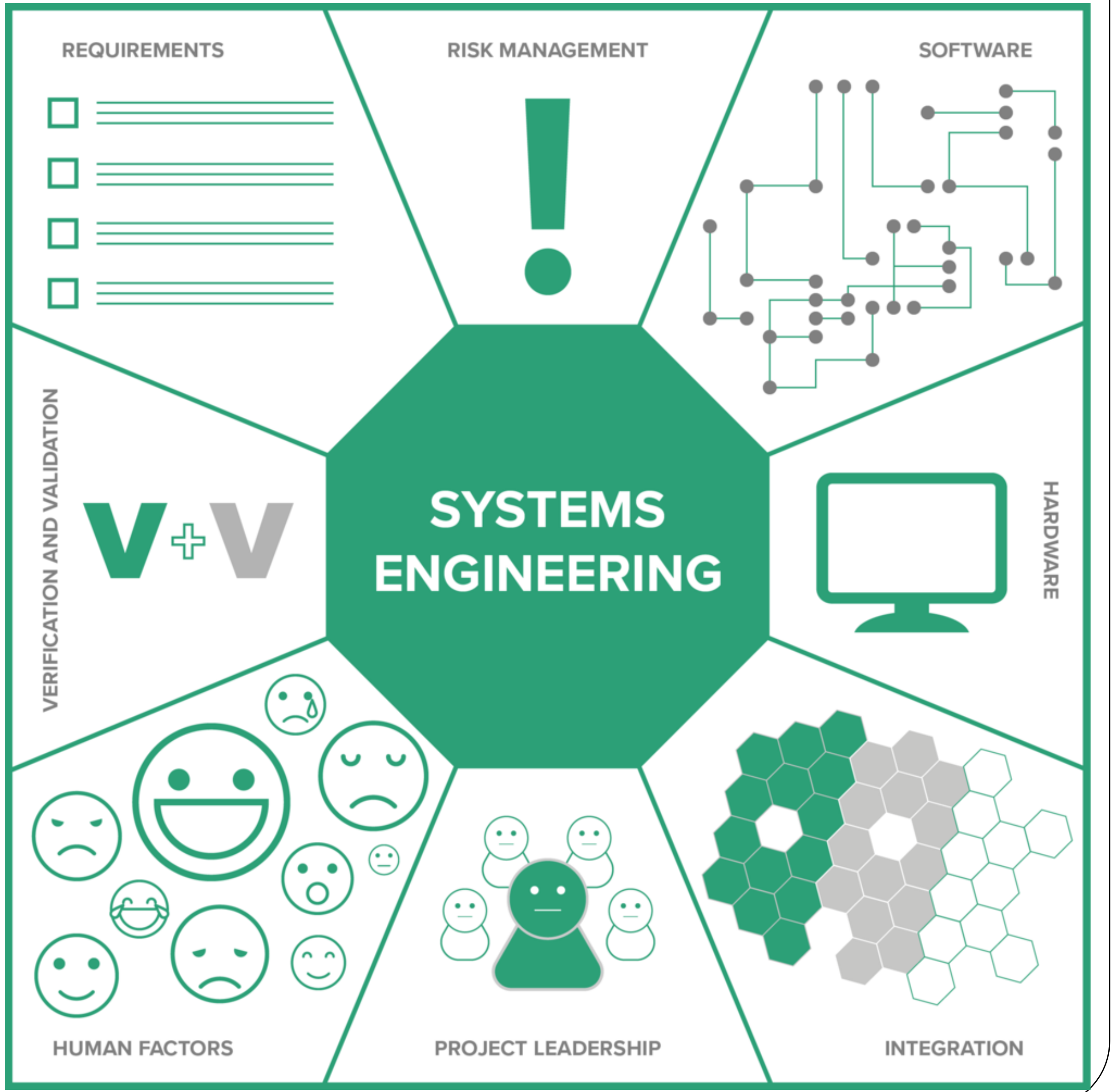
- Systems engineering helps in complex projects: spacecraft design, computer chip design, robotics, software integration, and bridge building.
- ➢ Modeling,
- ➢ Simulation,
- ➢ Requirements analysis,
- ➢ Scheduling,
- ➢ Manage complexity



https://en.wikipedia.org/wiki/Systems_engineering

Systems Engineering Activities

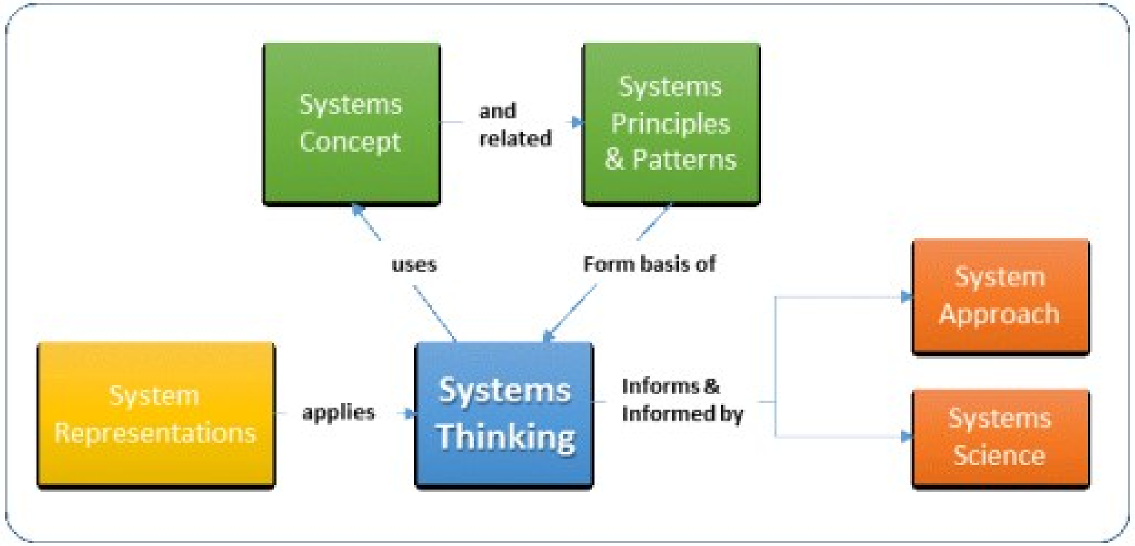
- Use of tools to validate assumptions or theories on systems and its internal interactions between components
  - System architecture,
  - System model, Modeling, and Simulation,
  - Optimization,
  - System dynamics,
  - Systems analysis,
  - Statistical analysis,
  - Reliability analysis, and
  - Decision making



[4] *Systems Engineering Fundamentals.* Defense Acquisition University Press, 2001

Systems Thinking

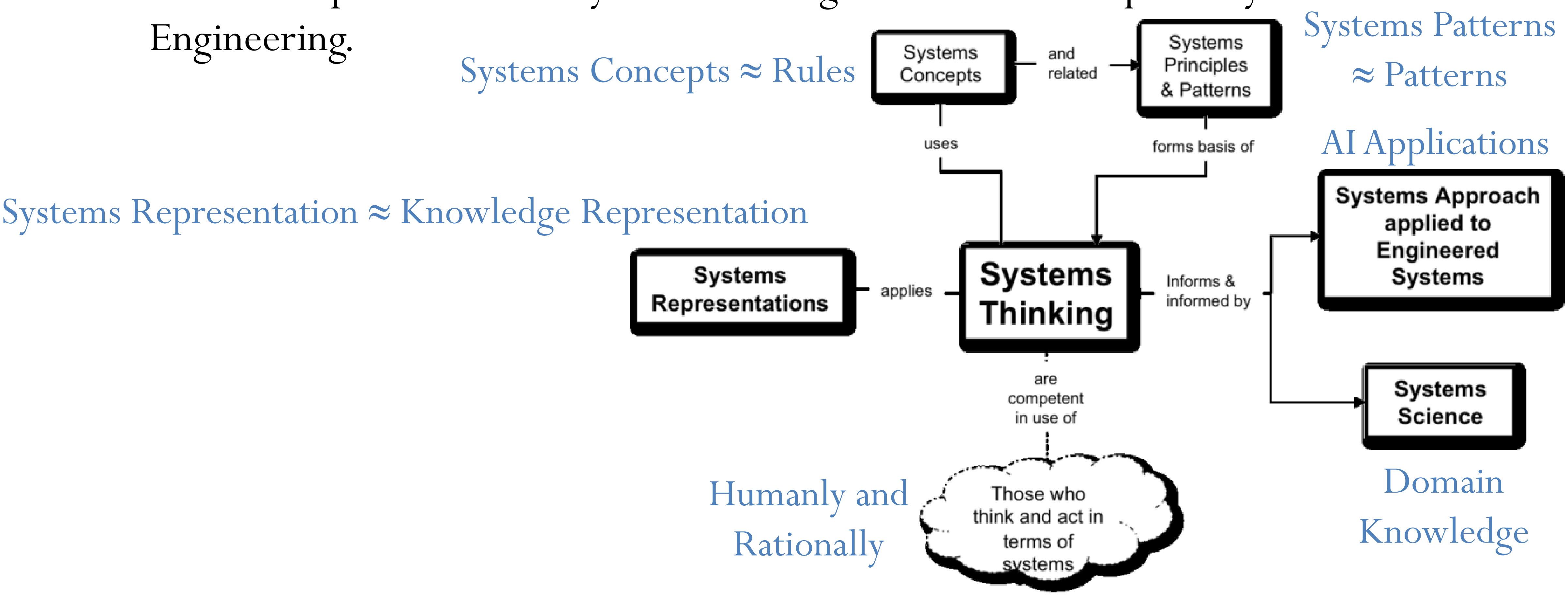
- Focuses and emphasizes on
  - behavior of the whole rather than the individual parts.
  - the interfaces between/among the subsystems.
- Ability to think about interactions between components of a system and their effect on functionalities.
- Understanding or Intervening in problem situations, based on the principles and concepts of the systems paradigm.



[8] “Systems Engineering Jr. Handbook”, *INCOSE LA Chapter and The Creativita Institute*.

Systems Thinking and Data Science

- Relationship between the System Thinking with other concepts of Systems Engineering.



[9] Cloutier, R. J. "The Guide to the Systems Engineering Body of Knowledge (SEBoK); v. 2.2." *INCOSE and The Trustees of the Stevens Institute of Technology: Hoboken, NJ, USA* (2016).

Data Science for Complexity, Chaos, Regularity

- **Self-adaptive systems:** Process of attention and **adaptation** to ensure appropriately identified boundaries, dependencies, and relationships.
- **Community detection:** Components has **cohesion**, or "**togetherness**"
- **Machine learning** or **Data mining** to calculate complexity, chaos, managing interdependency, and understanding choice.
  - More complex and chaotic because of inadequate concepts to explain.
  - Understanding reduce chaotic or complex.
- **Automation** for Similarities and Differences. E.g. Github code differencing
- **Machine learning** or **Data mining** to find Regularity
  - Regularity is a uniformity or similarity that exists in multiple entities or at multiple times.
  - Regularity in both natural systems and engineered systems.

[9] Cloutier, R. J. "The Guide to the Systems Engineering Body of Knowledge (SEBoK); v. 2.2." *INCOSE and The Trustees of the Stevens Institute of Technology: Hoboken, NJ, USA* (2016).

State Variables, Stability, and Determinism

- **Data Analytics** to study state variables, stability, and determinism of a system state
- A stable state is one in which a system will remain until another event occurs.
- State can be monitored using state variables, values of attributes.
- Set of possible values of state variables over time is called the "state space".
- State variables are continuous;
  - modeled using a finite state model (or state machine).
- A system can react, respond, or act.
- Stable system has stable states within an environment for a range of possible events.
- Deterministic systems have a one-to-one mapping of state variables to state space,
  - allowing future states to be predicted from past states.
- Non-Deterministic systems have a many-to-many mapping of state variables;
  - future states cannot be reliably predicted

[9] Cloutier, R. J. "The Guide to the Systems Engineering Body of Knowledge (SEBoK); v. 2.2." *INCOSE and The Trustees of the Stevens Institute of Technology: Hoboken, NJ, USA* (2016).

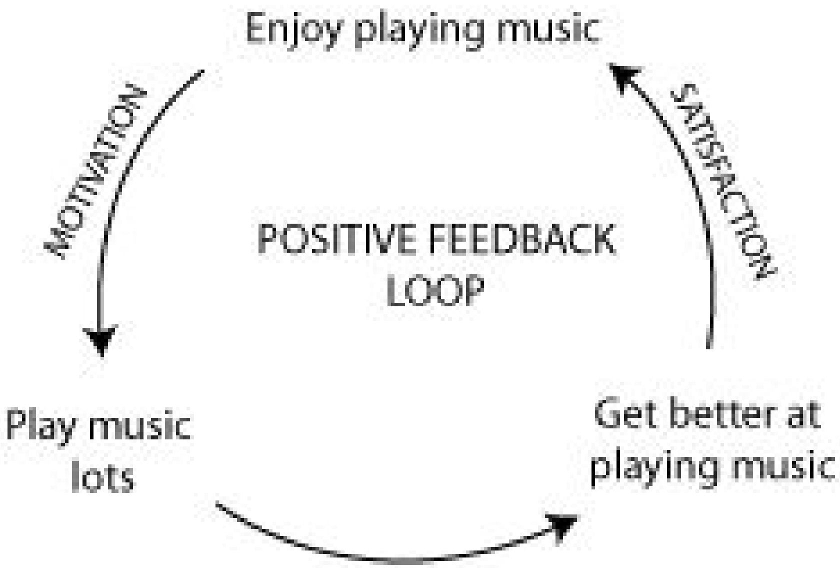
Systems attributes & environment

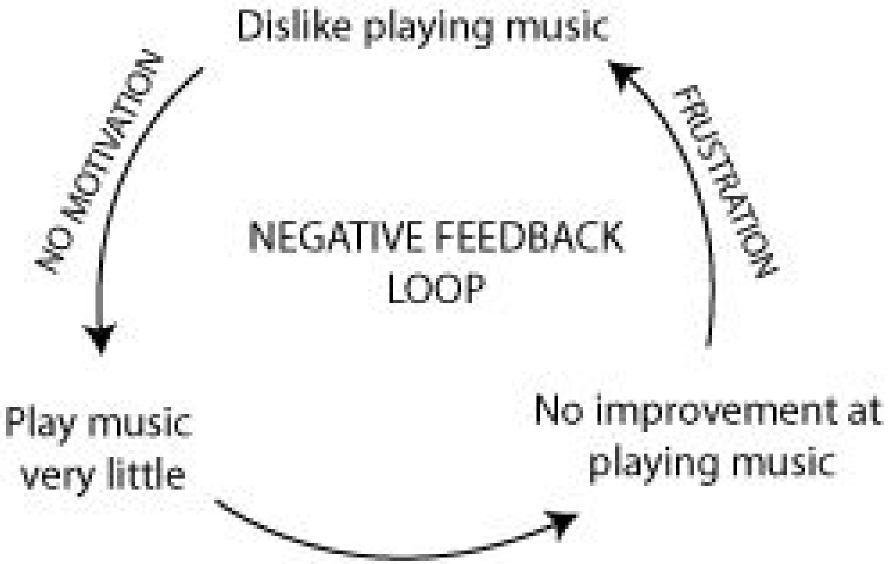
- Any quality or property of a system element is called an **attribute**.
- Attributes are used for **Data Analytics: Machine learning** and **Data mining**.
  - The state of a system is a set of system attributes at a given time.
  - A system event describes any change to the environment of a system, hence its state.
- Type of system based on attributes
  - Static - A single state exists with no events.
  - Dynamic - Multiple possible stable states exist.
  - Homeostatic - System is static but its elements are dynamic. The system maintains its state by internal adjustments.

[9] Cloutier, R. J. "The Guide to the Systems Engineering Body of Knowledge (SEBoK); v. 2.2." *INCOSE and The Trustees of the Stevens Institute of Technology: Hoboken, NJ, USA* (2016).

Control Behaviour

- System has interconnected component.
- Boundaries: Outer components of a system in an environment.
- Feedback loop: Process to self-correct based on reactions from other systems in the environment.

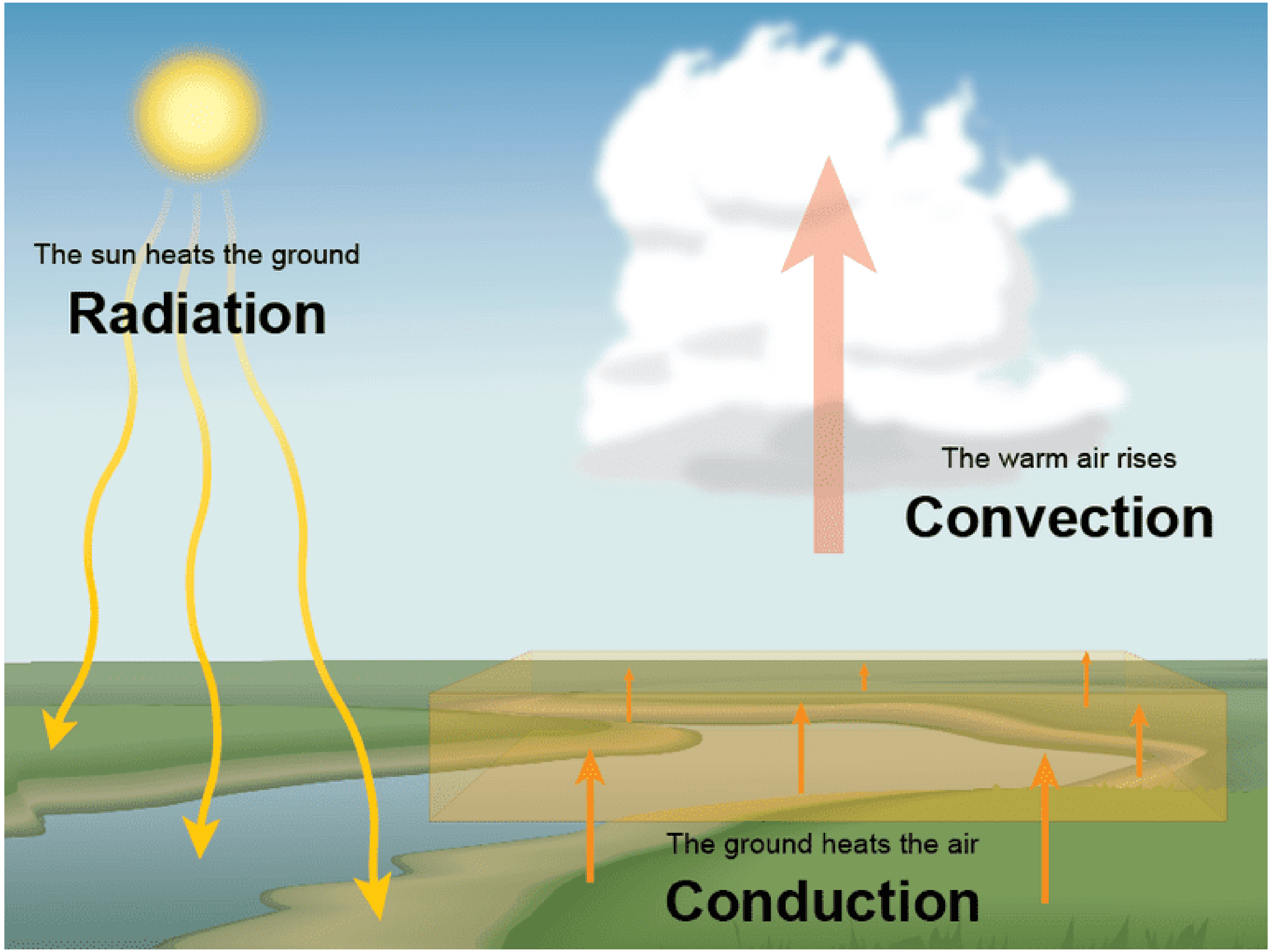




https://en.wikipedia.org/wiki/Systems_theory

Control Behaviour

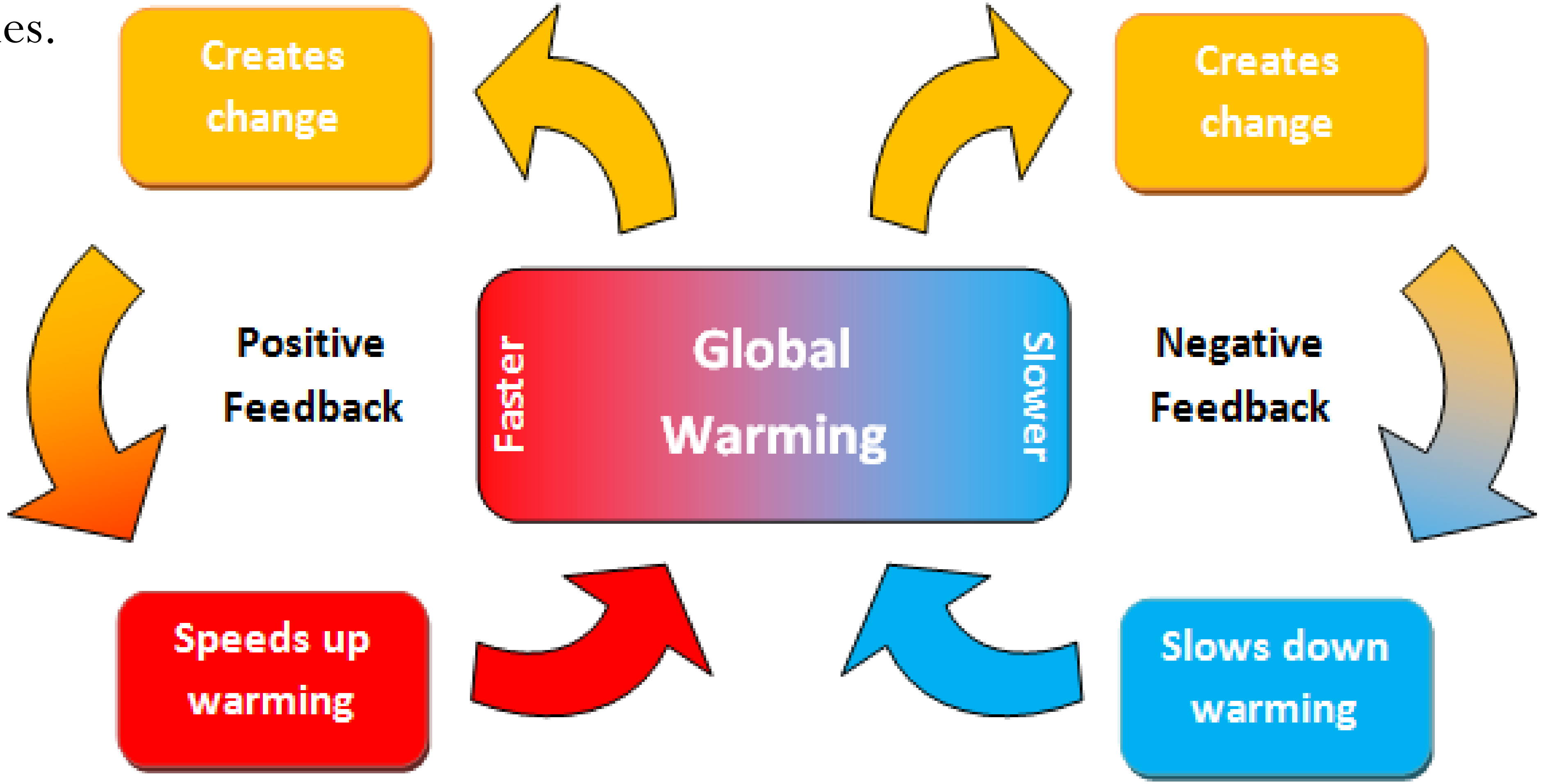
- Reciprocal transactions: Cyclical interactions such that systems influence one another.
- Throughput: Rate of energy transfer between a system and its environment over time.



https://en.wikipedia.org/wiki/Systems_theory

Control Behaviour

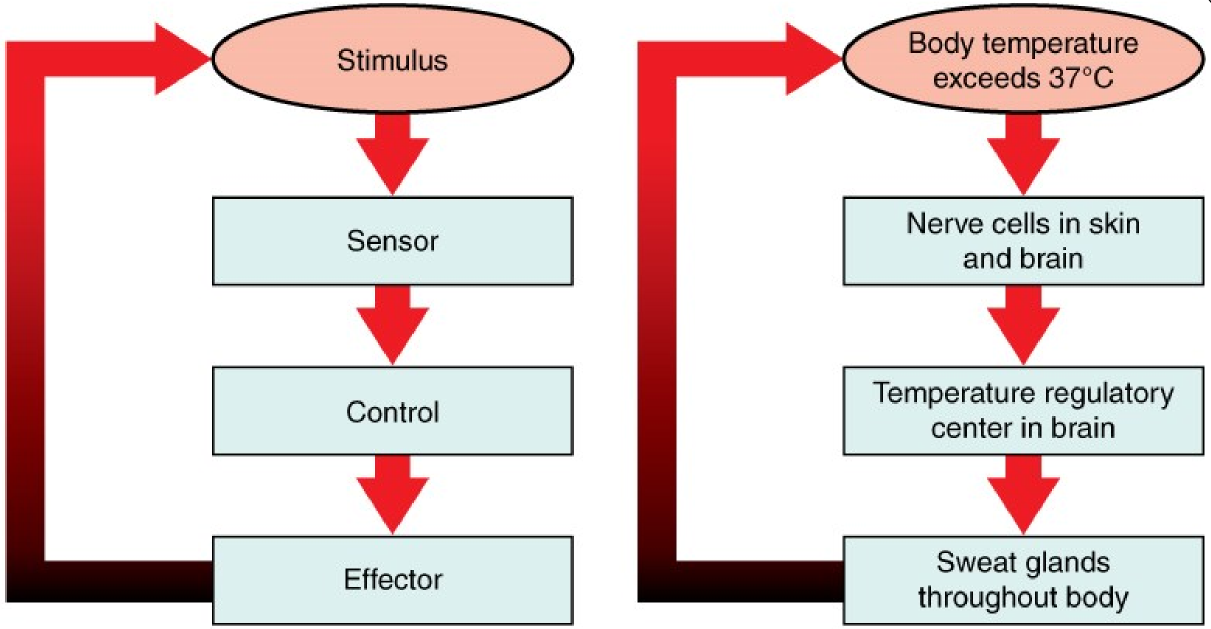
- Homeostasis: Tendency to be resilient w.r.t external disruption and to maintain functionalities.

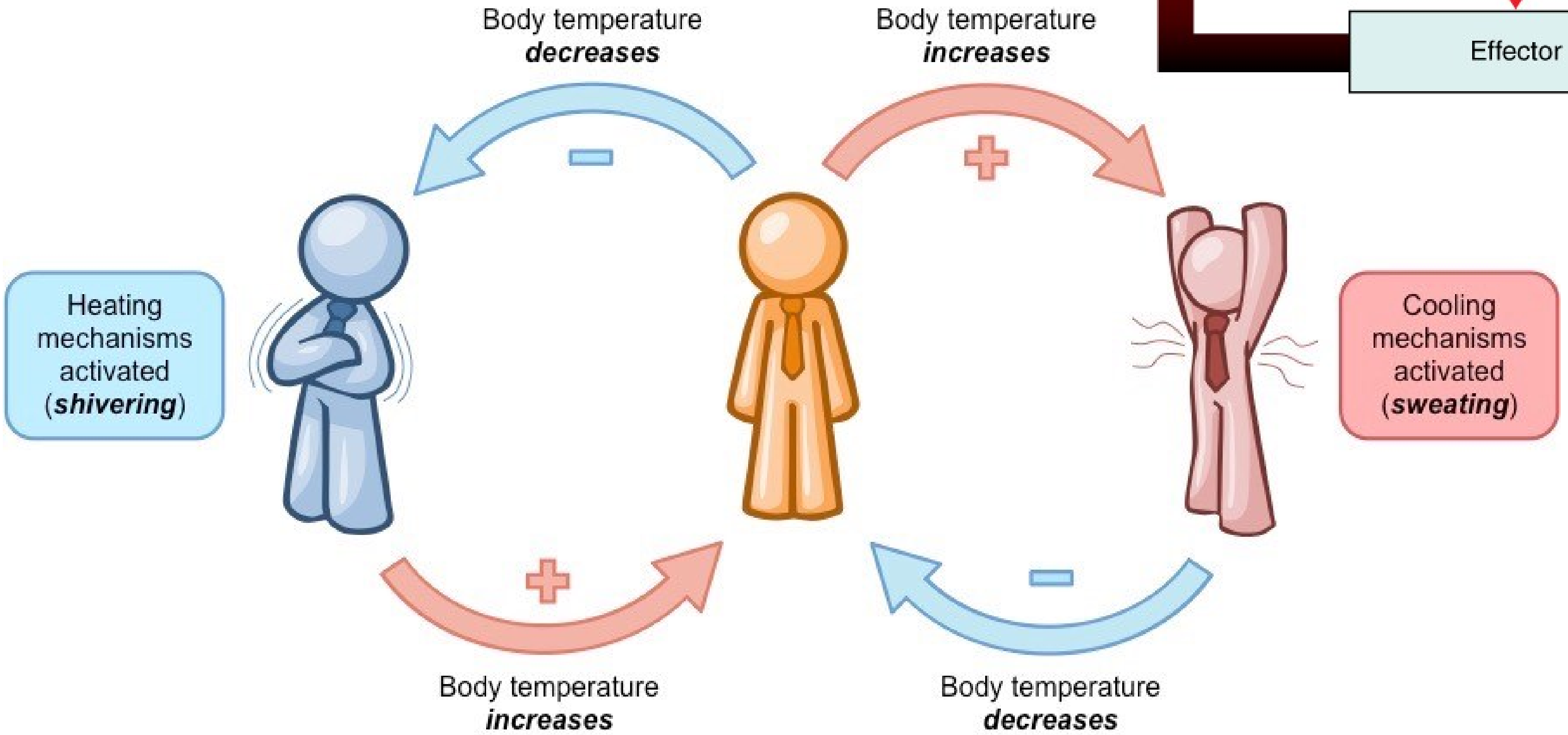


https://en.wikipedia.org/wiki/Systems_theory

Control Behaviour

- Adaptation: Tendency of making internal changes to protect itself and to maintain functionalities.

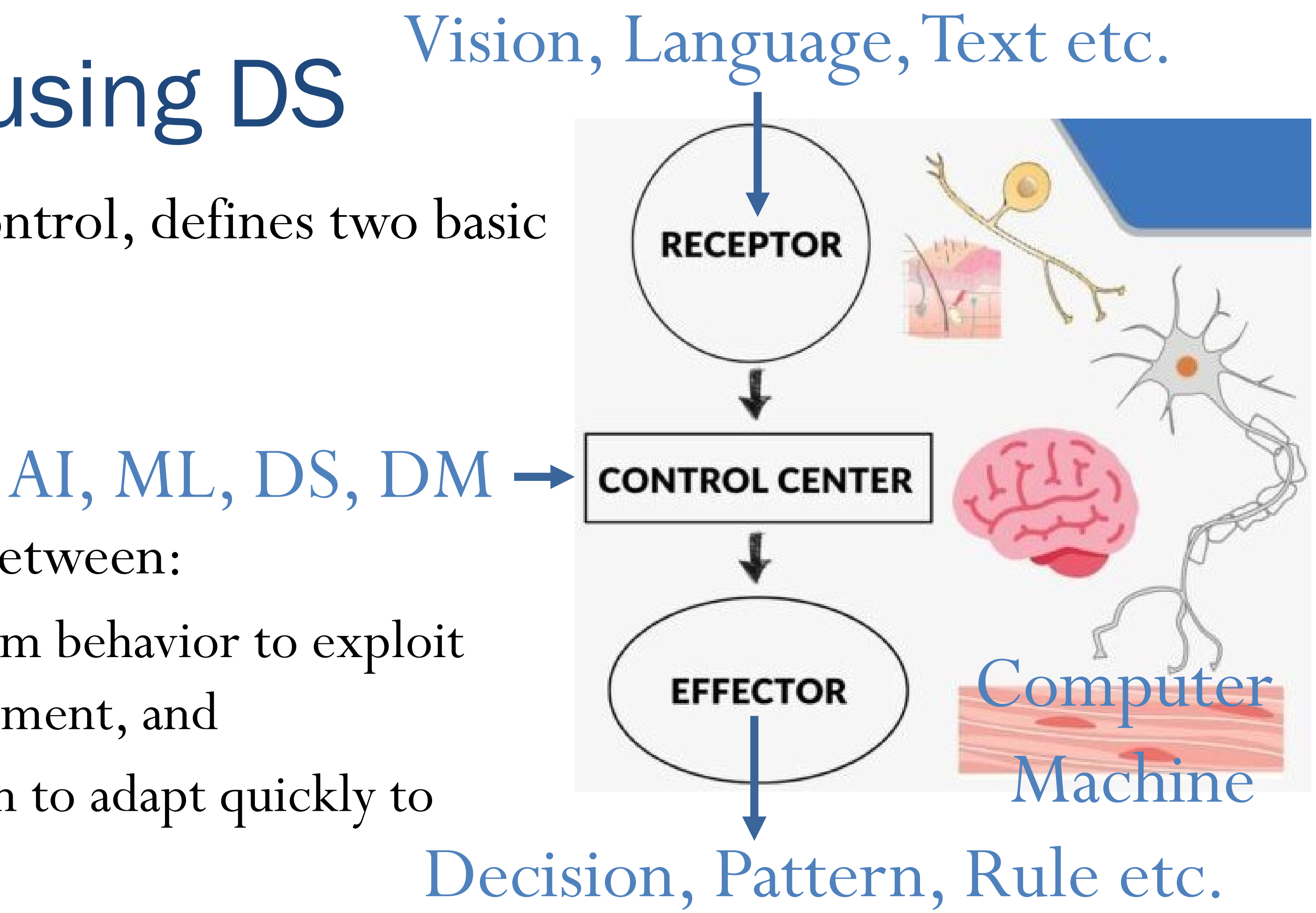




https://en.wikipedia.org/wiki/Systems_theory

Control Behaviour using DS

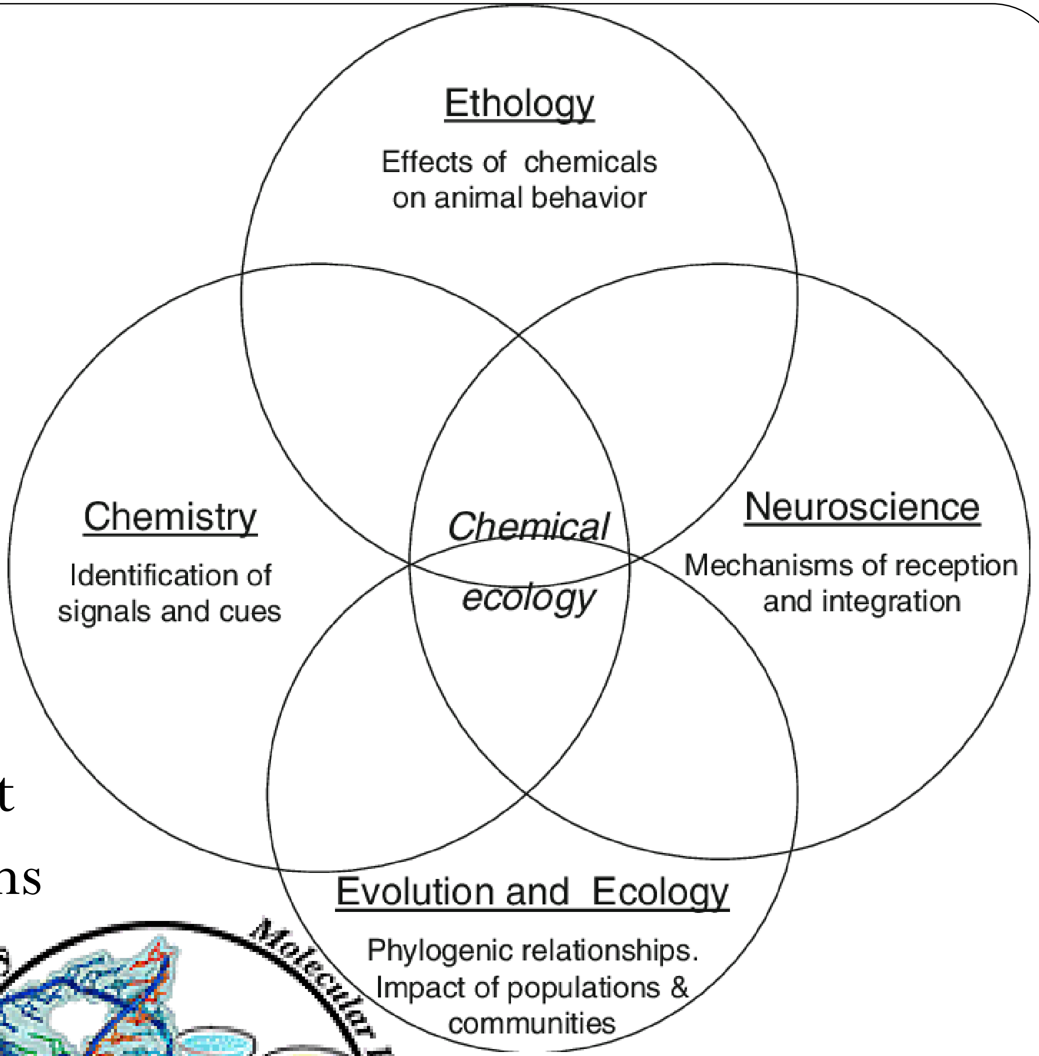
- **Cybernetics**, the science of control, defines two basic control mechanisms:
  - Negative feedback,
  - Positive feedback,
- **Control behavior** is a trade between:
  - Specialization, the focus of system behavior to exploit particular features of its environment, and
  - Flexibility, the ability of a system to adapt quickly to environmental change.

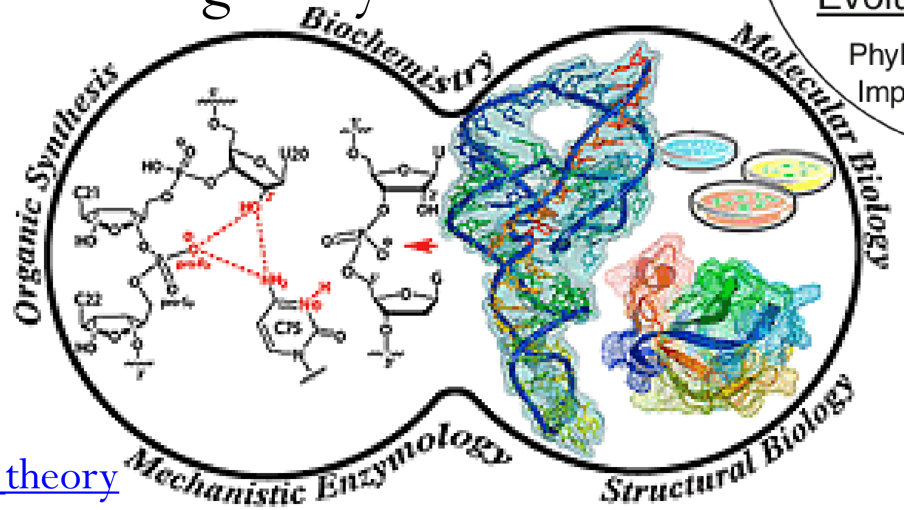


[9] Cloutier, R. J. "The Guide to the Systems Engineering Body of Knowledge (SEBoK); v. 2.2." *INCOSE and The Trustees of the Stevens Institute of Technology: Hoboken, NJ, USA* (2016).

Systems Examples

- **Systems biology** is bioscience research focusing on complex interactions in biological systems.
- **Systems chemistry** studies networks of interacting molecules.
- **Systems ecology** is a field of ecology (subset Earth science) that studies of ecological systems i.e. ecosystems.

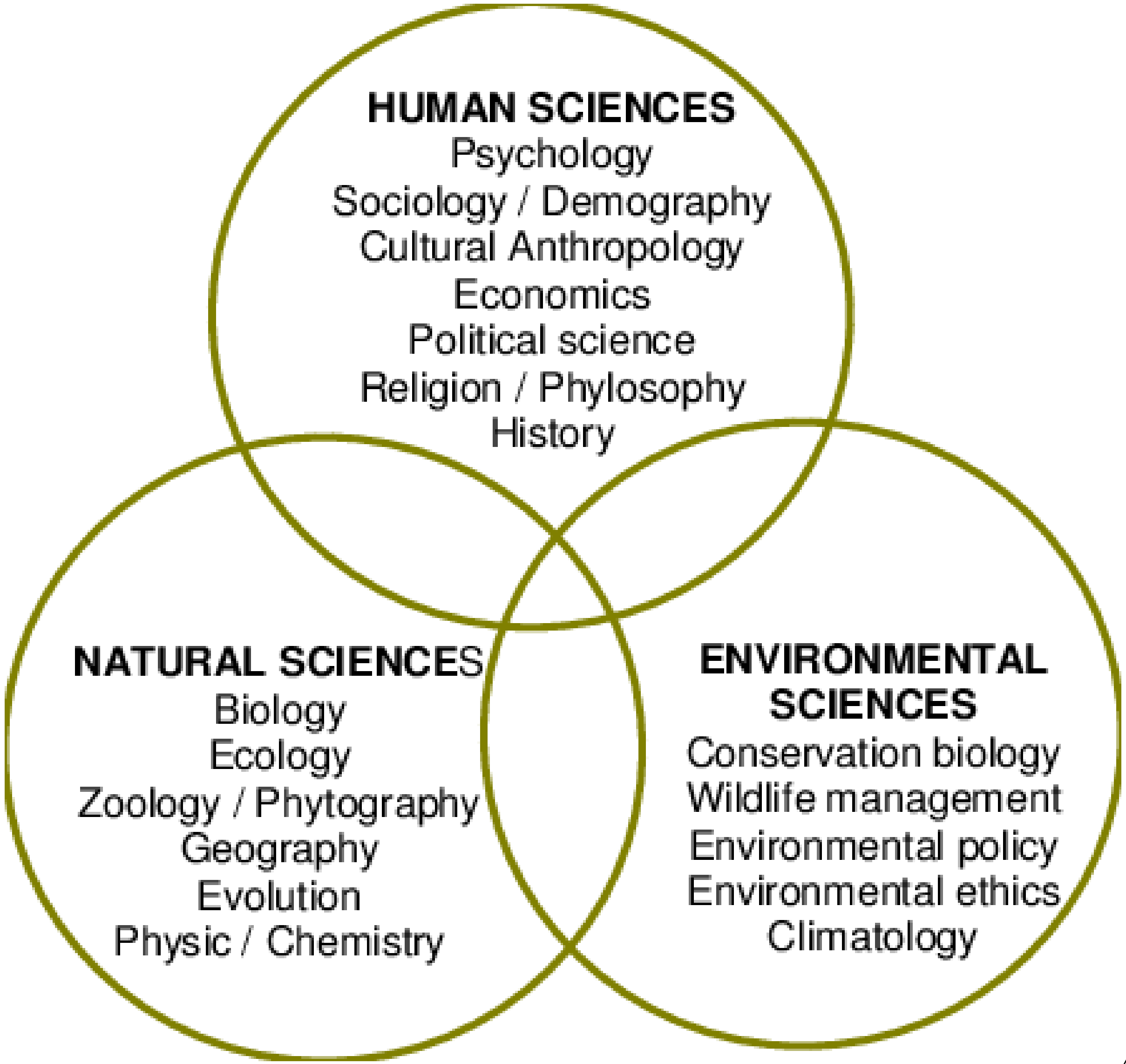




https://en.wikipedia.org/wiki/Systems_theory

Systems Examples

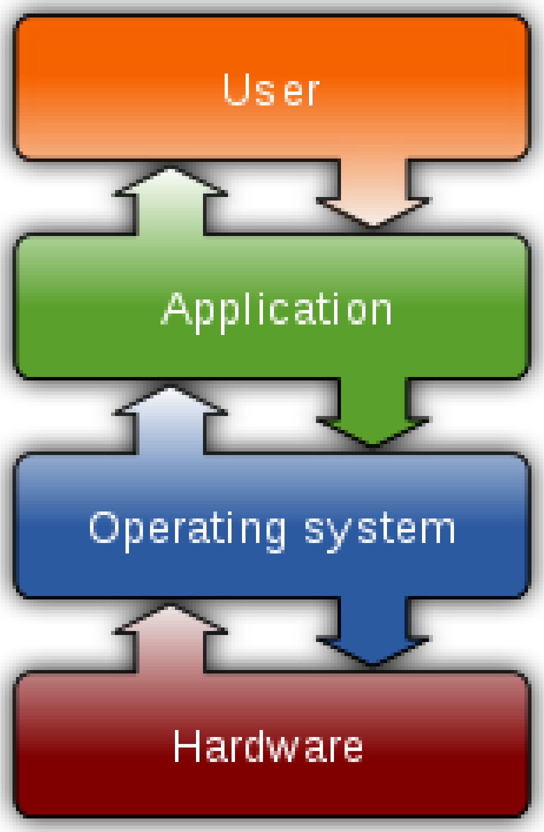
- **Systems psychology** studies human behavior and experience in complex systems.
- **Systems thinking** ability or skill to perform problem solving in complex system. With the Systems theory, the System Thinking can be learned.

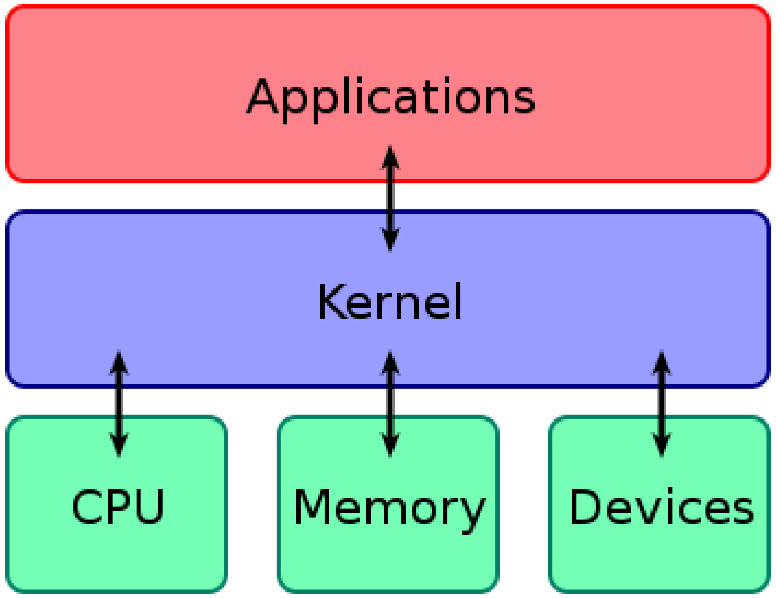


https://en.wikipedia.org/wiki/Systems_theory

Operating system

- Kernel manages the connection between application software with the hardware of a computing machine.

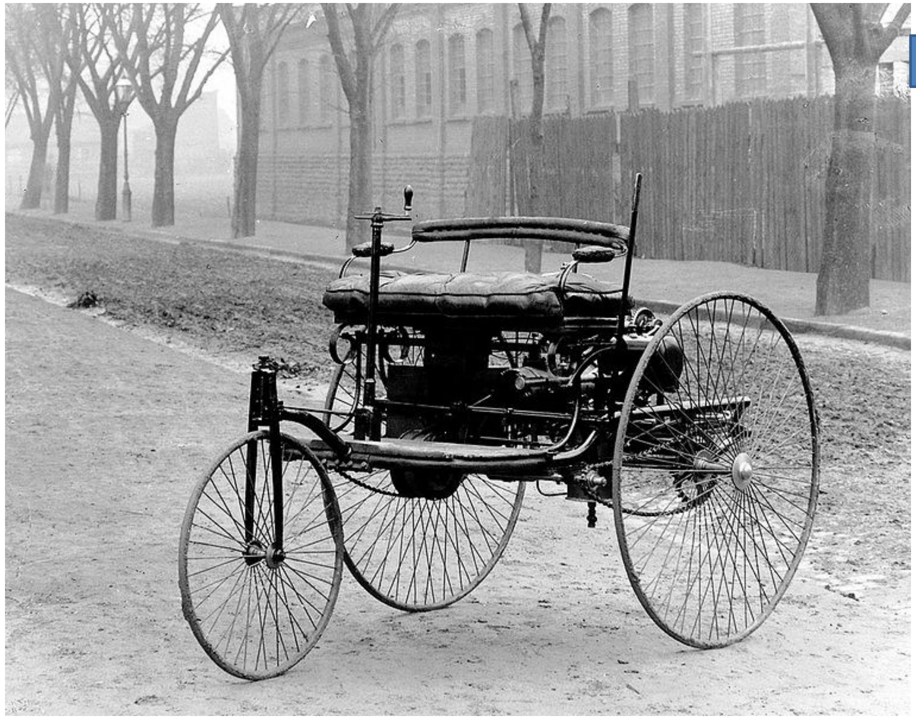




https://en.wikipedia.org/wiki/Operating_system

Automobile systems

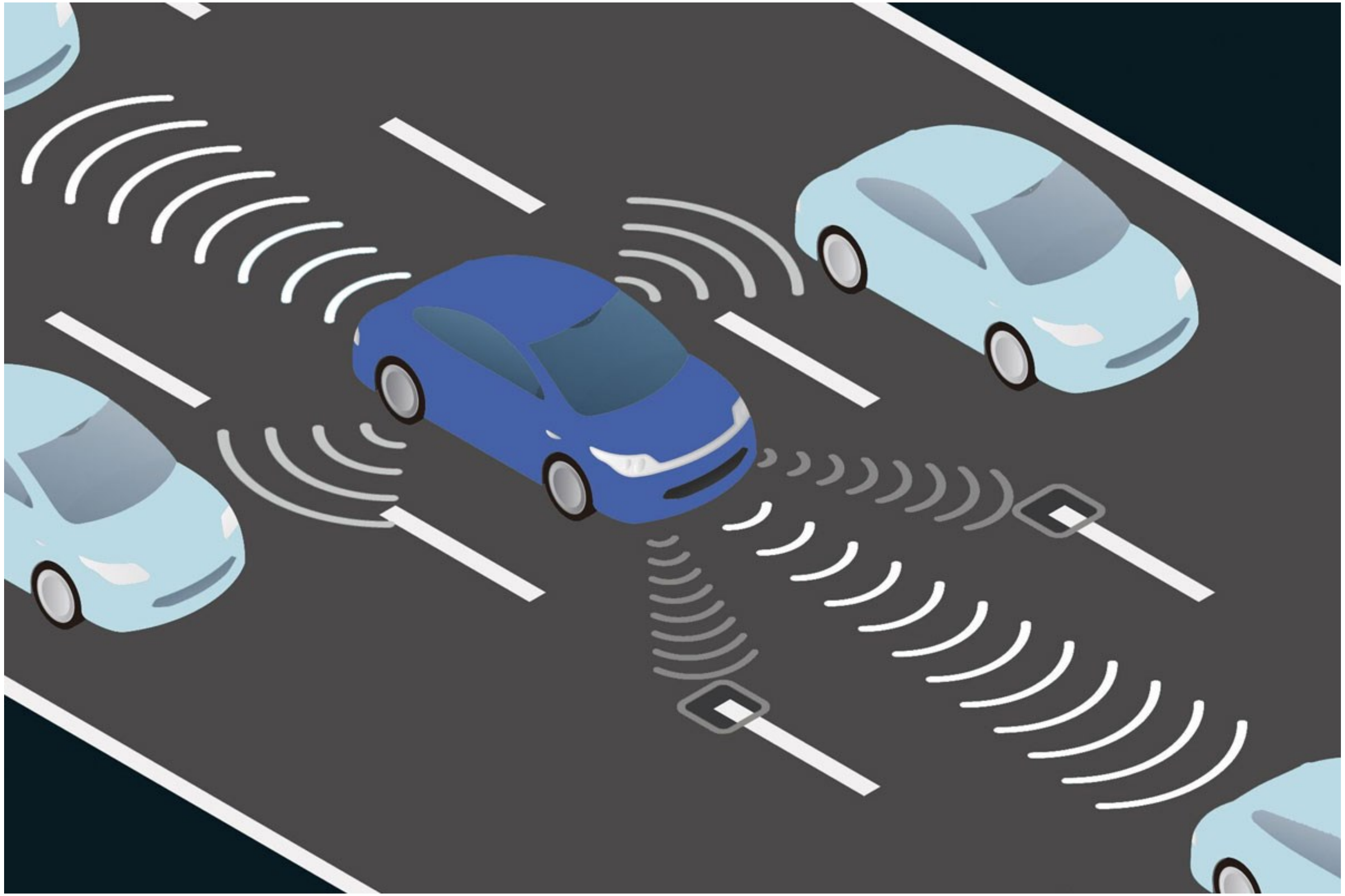
- Automobile systems from 1 engineer to 1000s of engineer working to built modern automobile



https://en.wikipedia.org/wiki/File:1885Benz.jpg

Autonomous Automobile Systems

Autonomous Automobile Systems



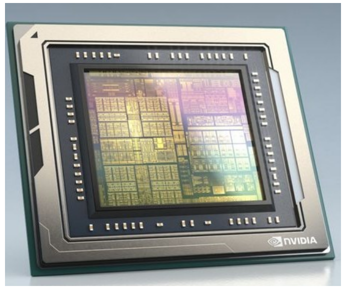
Autonomous Driving Cars

- aka. Self-driving car
- a vehicle that is capable of sensing its environment and moving safely with little or no human input.
- Self-driving cars combine a variety of sensors to perceive their surroundings, such as radar, sonar, GPS, odometry, and inertial measurement units.
- Advanced control systems interpret sensory information to identify appropriate navigation paths, as well as obstacles and relevant signage.

- https://www.youtube.com/watch?v=tlThdr3O5Qo
- https://www.youtube.com/watch?v=0GnysB0rO3s

https://en.wikipedia.org/wiki/Self-driving_car

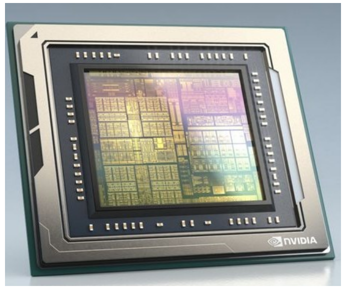
Autonomous Driving Cars



- NVIDIA DRIVE AGX
- Software-Defined Platform for Autonomous Machines
  - deep neural networks run simultaneously in autonomous vehicles and robots
- The platform is powered by a new system-on-a-chip (SoC) called Orin, which consists of 17 billion transistors and is the result of four years of R&D investment.
- The Orin SoC integrates NVIDIA's next-generation GPU architecture and Arm Hercules CPU cores.
- deliver 200 trillion operations per second — nearly 7x the performance of NVIDIA's previous generation Xavier SoC.
- Orin is designed to handle the large number of applications (e.g. deep learning and computer vision)

https://nvidianews.nvidia.com/news/nvidia-introduces-drive-agx-orin-advanced-software-defined-platform-for-autonomous-machines

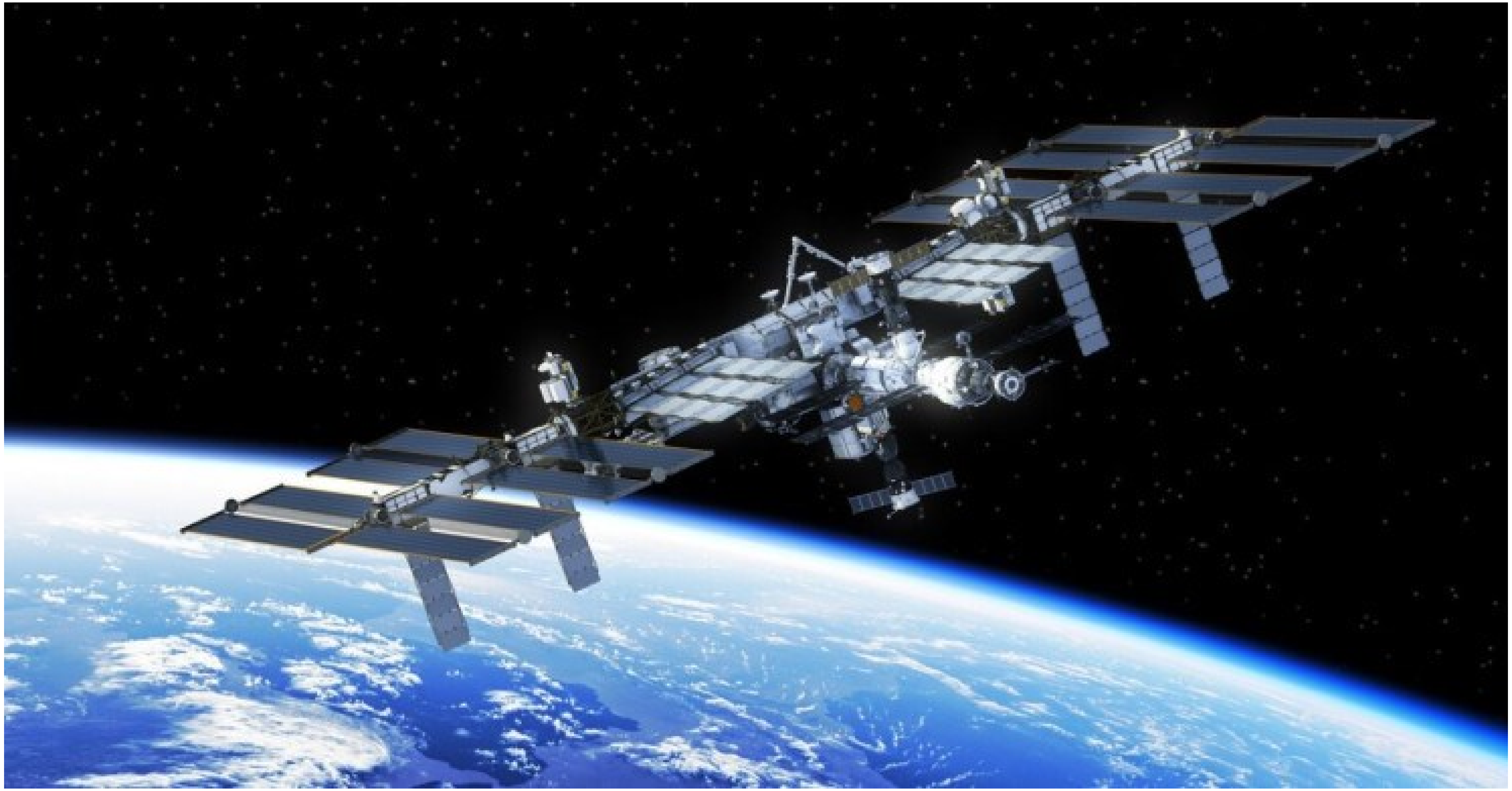
Autonomous Driving Cars



- NVIDIA DRIVE AGX
- Developed to enable architecturally compatible platforms that scale full self-driving vehicle, enabling to develop large-scale and complex families of software products.
- "Creating a safe autonomous vehicle is perhaps society's greatest computing challenge," said Jensen Huang, founder and CEO of NVIDIA.
- "The amount of investment required to deliver autonomous vehicles has grown exponentially, and the complexity of the task requires a scalable, programmable, software-defined AI platform like Orin."

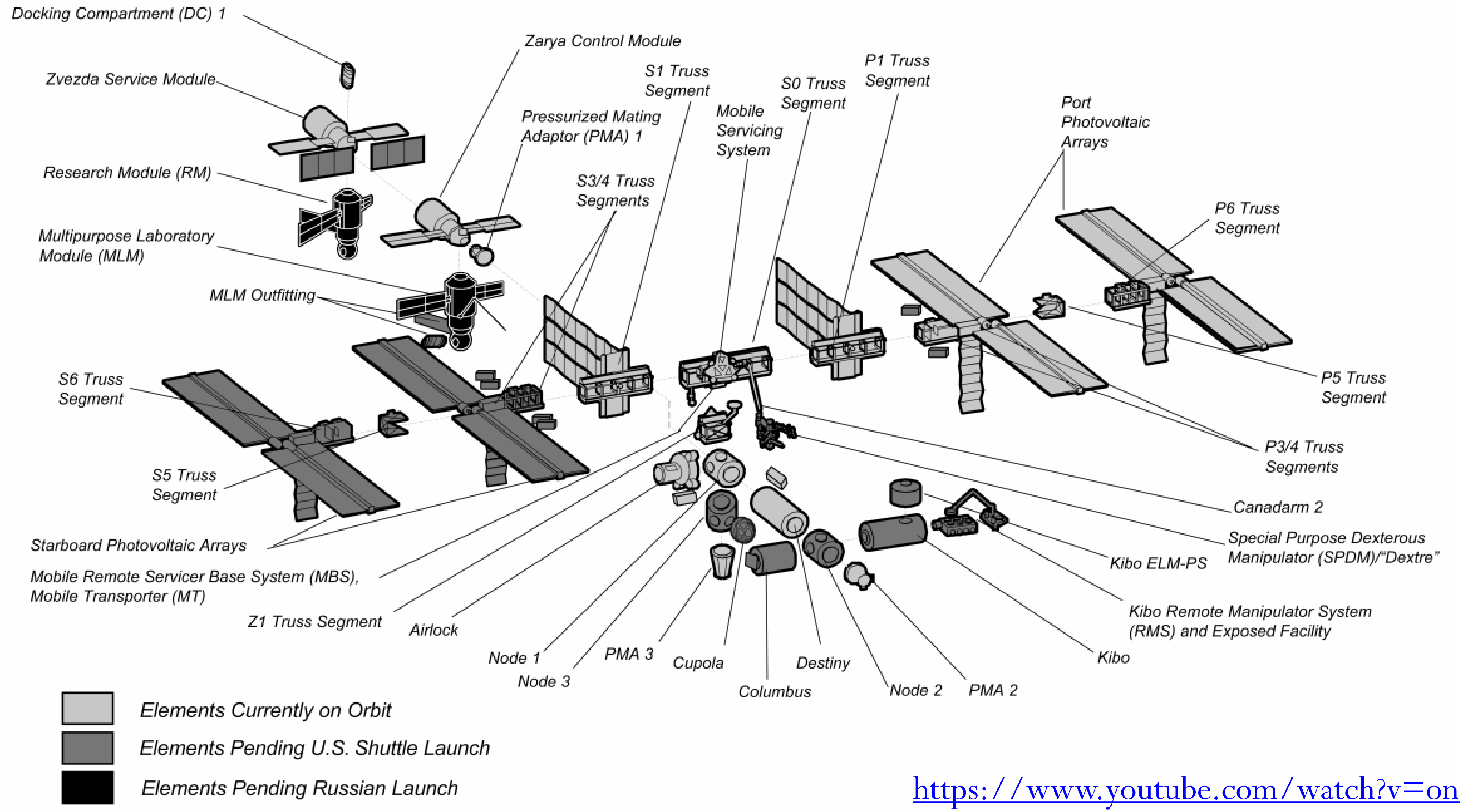
https://nvidianews.nvidia.com/news/nvidia-introduces-drive-agx-orin-advanced-software-defined-platform-for-autonomous-machines

International Space Station (ISS)



https://en.wikipedia.org/wiki/Systems_engineering

ISS Components

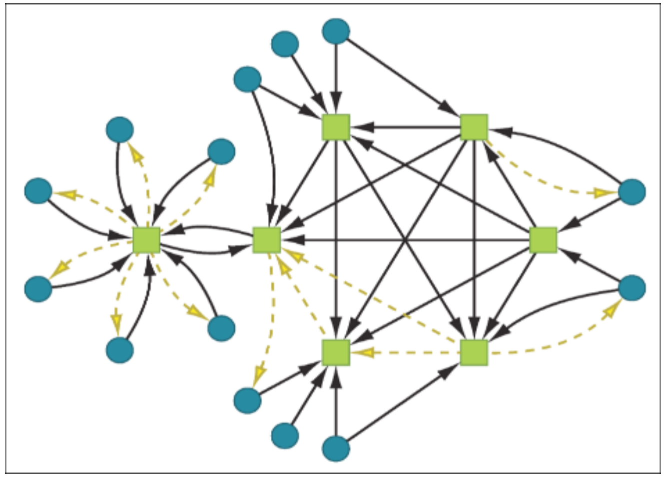


https://www.youtube.com/watch?v=onVhbeY7nLM

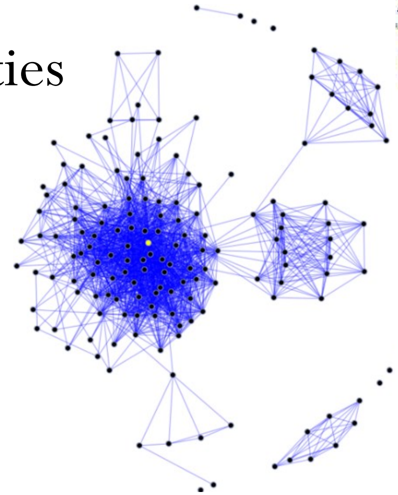
[7] *International Space Station Basics*, National Aeronautics and Space Administration (NASA).

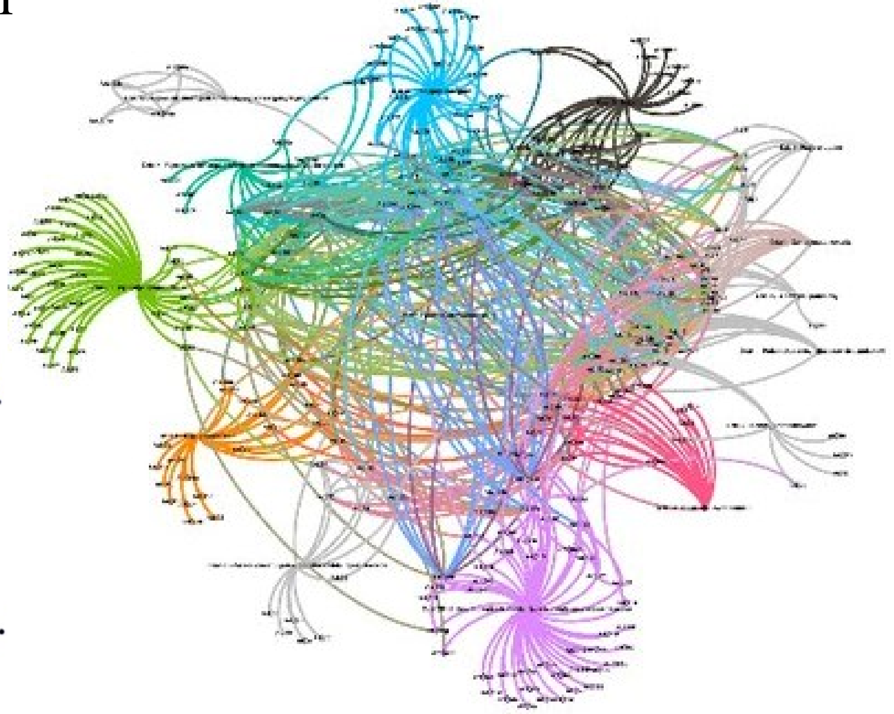
System Network – Data

Systems and Graph Theory



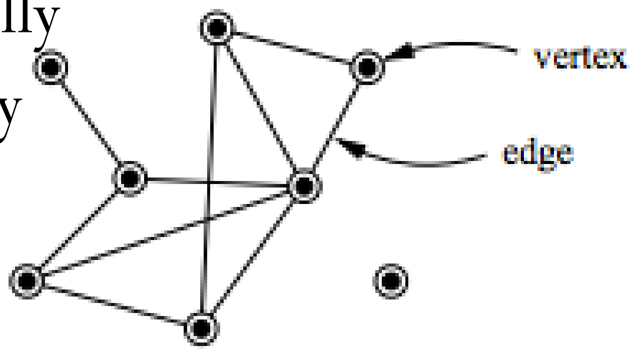
- **Data represented with a network (graph)**
  - nodes represent the components (or entities) and
  - edges (or links) represents entities' interactions.
- Depicts collection of discrete objects and relationships between them
  - persons within an organization,
  - logic gates in a circuit,
  - genes in gene regulatory networks, or
  - between any other set of related entities





Graph Theory

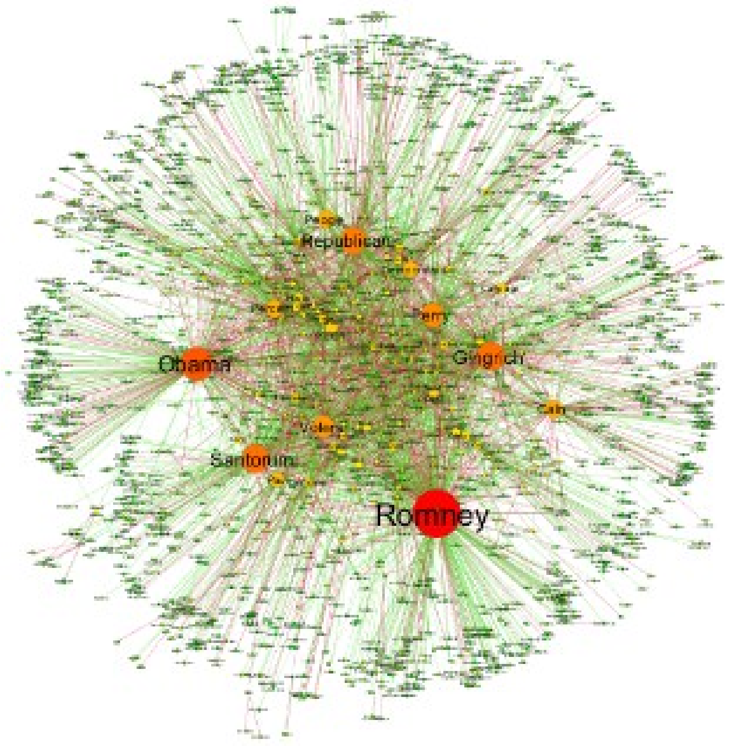
- **Data represented with a network (graph)**
- Study of mathematical structures used to model relationship between discrete elements of a set
- Vertices (or nodes) are connected by edges (or links)
- Undirected graphs, where edges link two vertices symmetrically
- Directed graphs, where edges link two vertices asymmetrically
- Graph is an ordered triple $G = (V, E, \emptyset)$
  - V is a set of vertices (or nodes)
  - E is a set of edges (or links)
  - $\emptyset: E \rightarrow \{\{x, y\} \mid x, y \in V \text{ and } x \neq y\}$ an incidence function mapping every edge to an unordered pair of vertices



Network Theory

- **Data represented with a network (graph)**
- Graph representing symmetric or asymmetric relations between discrete objects
- Graph with nodes and/or edges have attributes (e.g. names)
- Network theory is a subdomain of Graph theory
- Network theory is applied graph theory

Graph or Network Theories



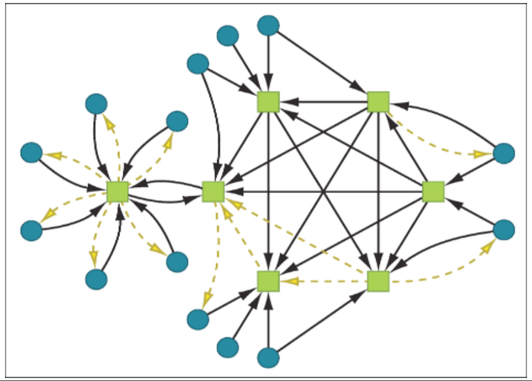
- **System Data properties (patterns)**
- **Graph colouring:** e.g. coloring a graph so that
  - no two adjacent vertices have the same color, or
  - no two coincident edges are the same color
- **Matching** is a set of edges without common vertices
- **Route problems:** e.g. shortest path problem
  - minimum spanning tree (MST): a subset of the connected edges that connects all the vertices together, without any cycles, and with minimum possible total edge weight.
  - traveling salesman problem (TSP): "Given a list of cities and the distances between each pair of cities, what is the shortest possible route that visits each city exactly once and returns to the origin city?"

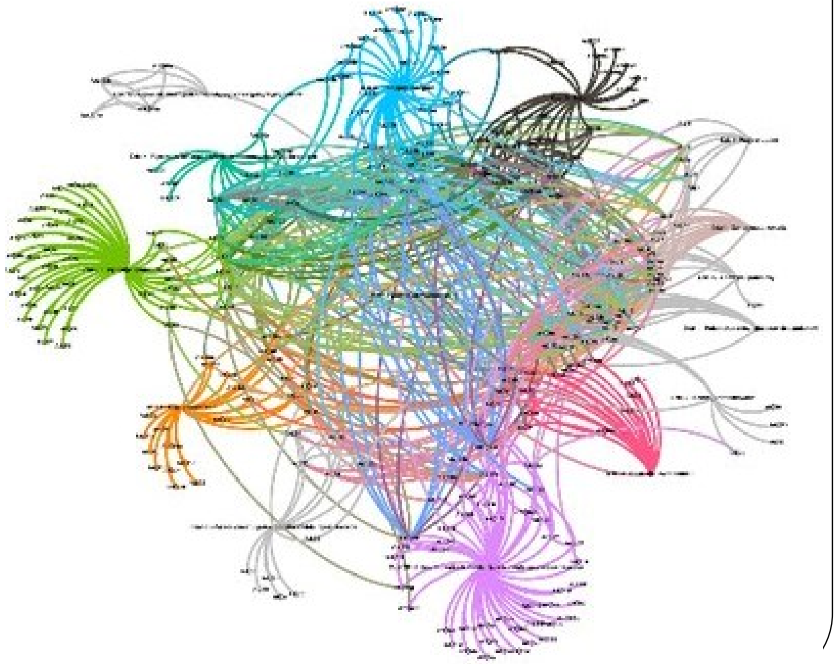
Graph or Network Theories

- **System Data properties (patterns)**
- **Network flow:** "Directed graph where each edge has a capacity and each edge receives a flow, where the amount of flow on an edge cannot exceed the capacity of the edge"
- **Transport problem:** study of optimal transportation and allocation of resources.
- **Trans-shipment problem:** a subgroup of transportation problems, where, transportation may or must go through intermediate nodes, possibly changing modes of transport
- **Critical path analysis:** identifying the longest dependent activities and measuring the time required to complete a project
- **PERT** to analyze and represent the tasks involved in completing a given project

System Data representation

- **Data Represented with a network (graph)**
  - nodes represent the components (or entities) and
  - edges (or links) represents entities' interactions.
- Depicts collection of discrete objects and relationships between them
  - persons within an organization,
  - logic gates in a circuit,
  - genes in gene regulatory networks, or
  - between any other set of related entities



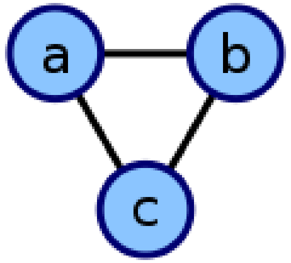


System Data representation

- **System Data representation with network (graph)**
- Complex systems can be represented by Complex networks
- Complex networks can represent connections between component (or entities)
- Complex network can be represented as a graph with
  - vertices (or nodes) representing entities and
  - edges (or links, or connections) representing the relationship between components.

System Data representation

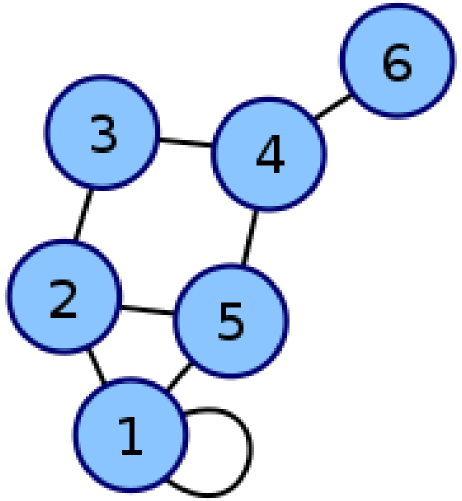
- **System Data representation with network (graph)**
- Adjacency list is a collection of unordered lists used to represent a finite graph.
- List contains the set of neighbors of a vertex in the graph.
- Undirected graphs:
  - data structure two different linked list nodes for each edge



Adjacency list representation	
a	b
a	c
b	a
b	c
c	a
c	b

System Data representation

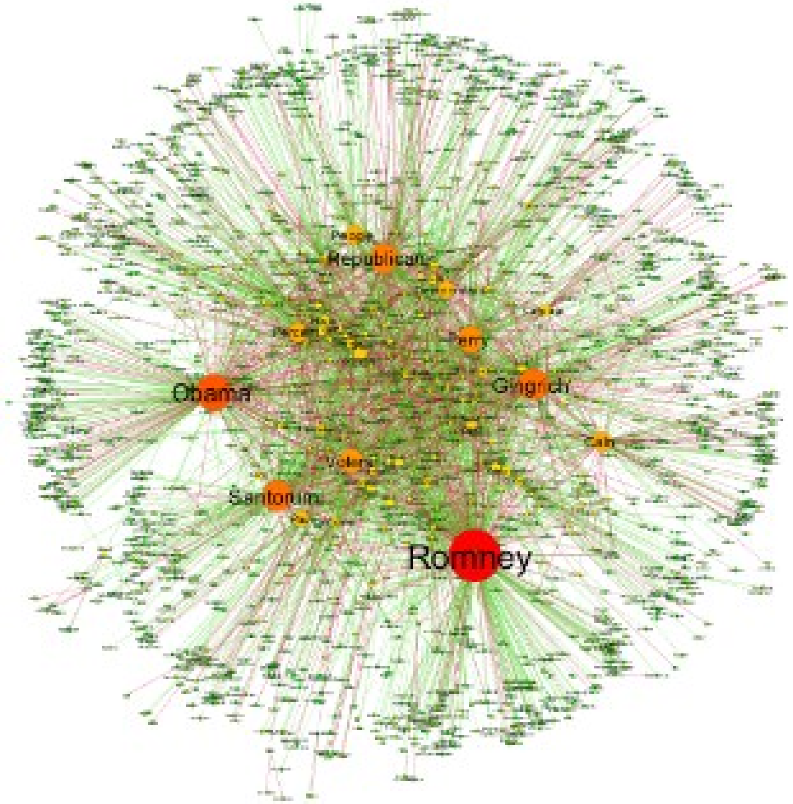
- **System Data representation with network (graph)**
- Adjacency matrix is a square matrix used to represent a finite graph.
- Matrix with rows and columns are indexed by vertices
- an adjacency list is more space-efficient than an adjacency matrix (stored as a two-dimensional array)
- adjacency list is as simple
- Matrix are 2-Dimensional representation.



$$\begin{pmatrix} 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}$$

System Data representation

- **System Data representation with Complex Networks**
- Non-trivial topological features in networks representing real systems
- e.g. theories: Network motifs are small subgraphs that are over-represented in the network
- Examples of Complex networks are
  - computer networks,
  - biological networks,
  - technological networks,
  - brain networks,
  - climate networks and
  - social networks.

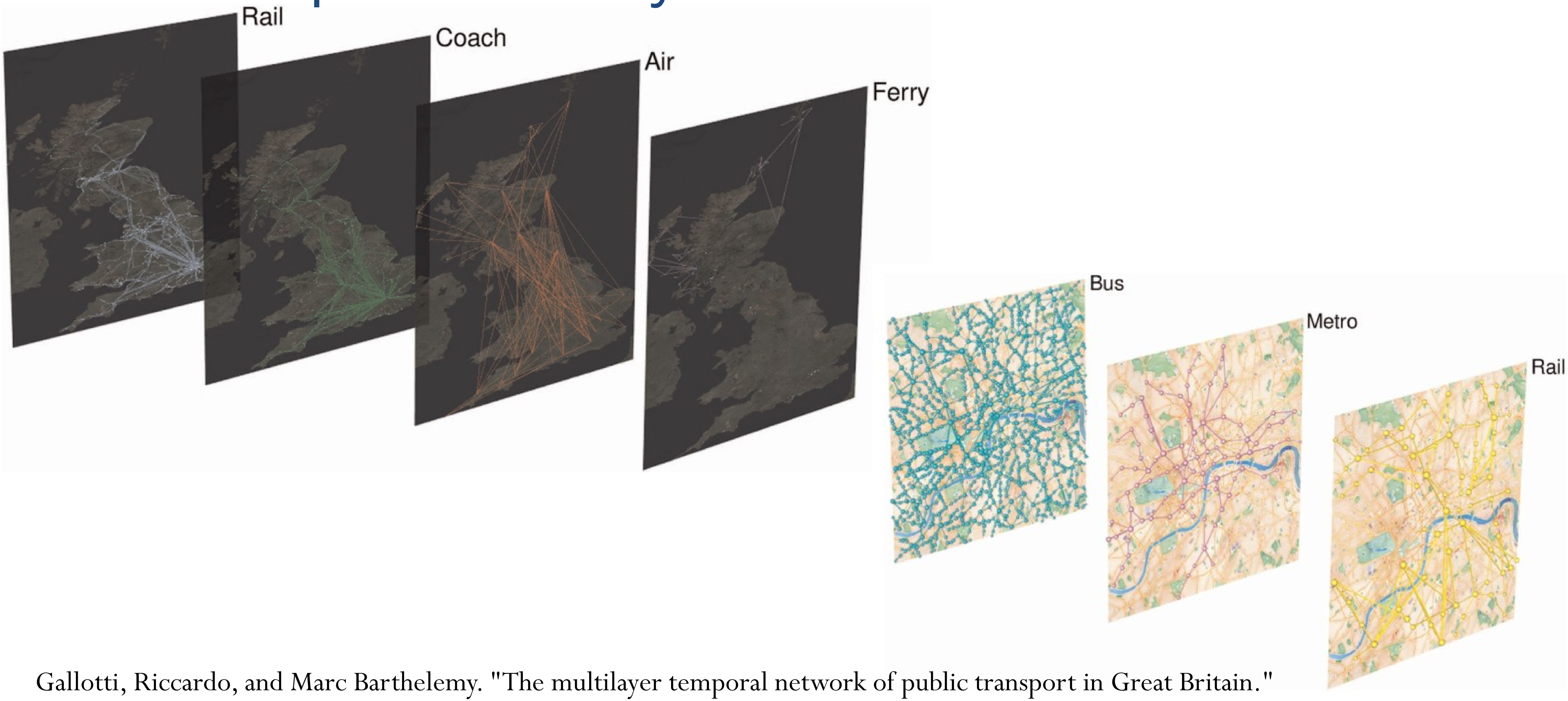


https://en.wikipedia.org/wiki/Complex_network

System Data representation

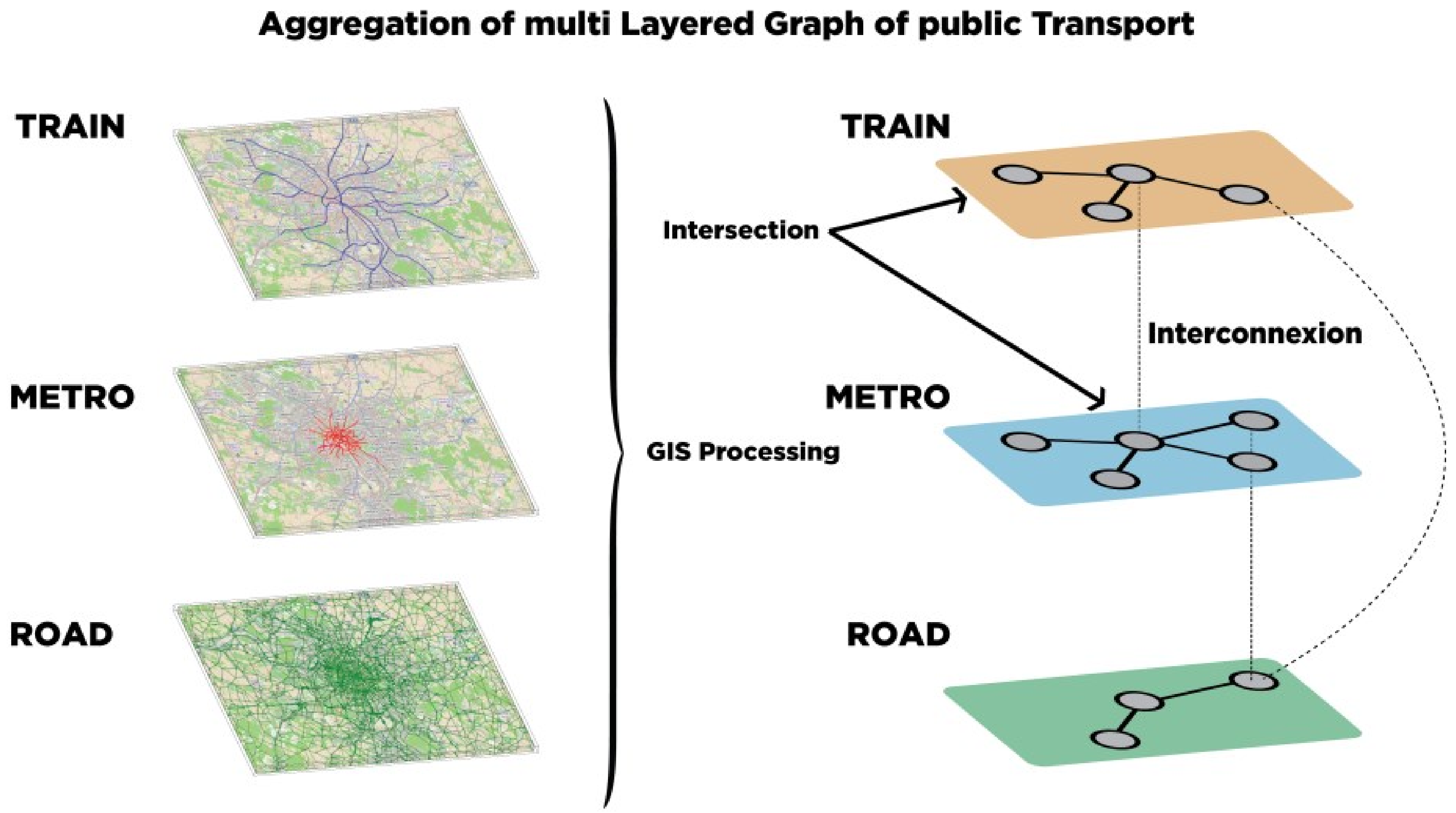
- **System Data representation with Multilayer Network representation:**
- Complex systems can be described by representing many interacting entities in
  - a Single network or
  - Multilayer networks
- Examples social networks, neural networks
- Co-occurring network structure made by links, activity as the nodes (e.g. community structure between multi-links).
- Dynamic entity relationships can be represented with multilayer networks, multiplex networks, and network of networks.

Transport multi-layer network



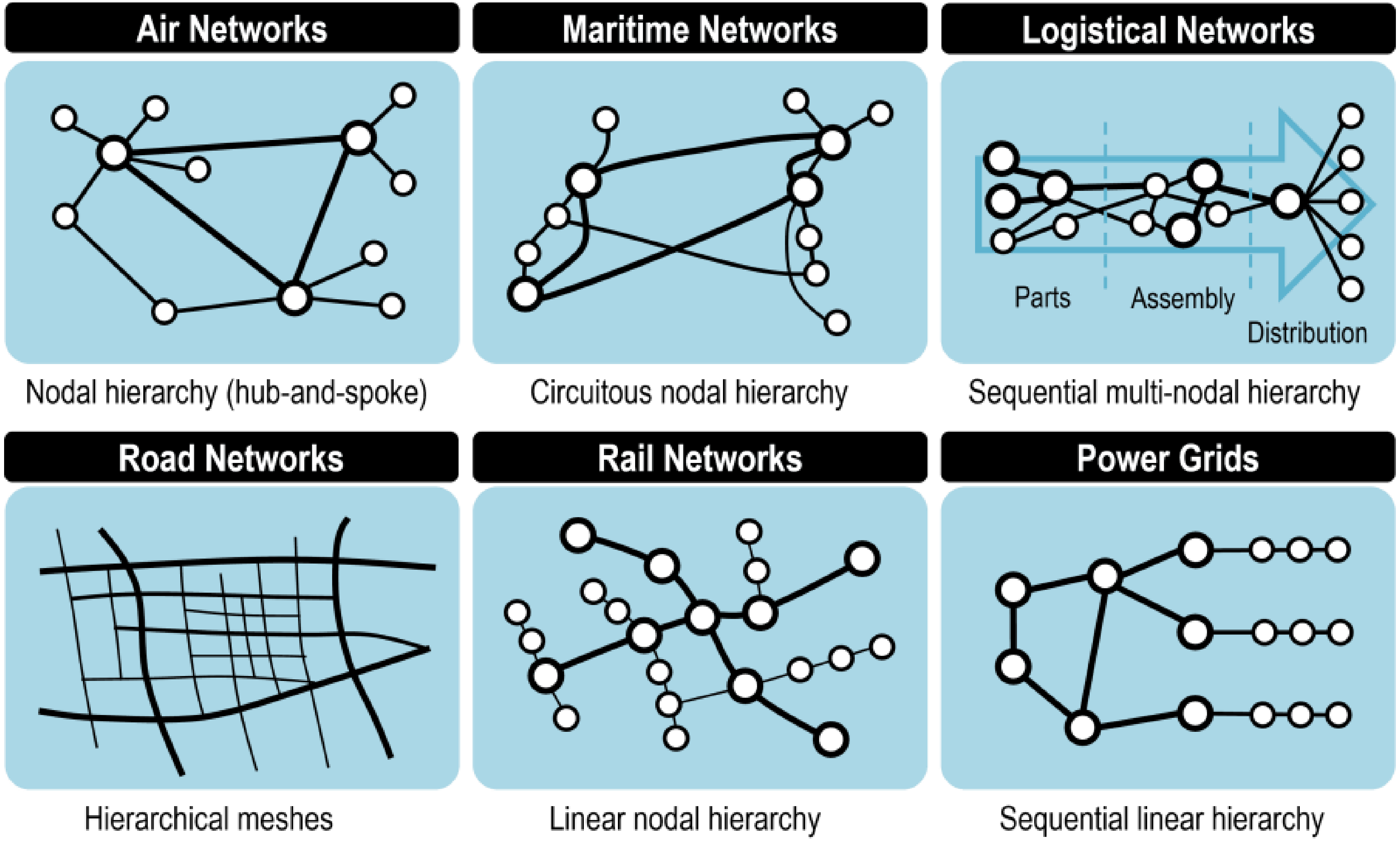
Gallotti, Riccardo, and Marc Barthelemy. "The multilayer temporal network of public transport in Great Britain." Scientific data 2.1 (2015): 1-8.

Transport multi-layer network



Asgari, Fereshteh, et al. "CT-Mapper: Mapping sparse multimodal cellular trajectories using a multilayer transportation network." Computer Communications 95 (2016): 69-81.

Types of Transportation Networks



https://transportgeography.org/contents/chapter9/transportation-and-disasters/transportation-networks-vulnerabilities/

System Data analysis

- **Electric power systems data analysis:**
  - a graph consists from representing electric power aspects (e.g., transmission line impedances)
- **Biological network data analysis:**
  - analysis of molecular networks
  - visualize the nature and strength of interactions between species
  - Gene Regulatory Networks (GRN), Metabolic networks, Protein-Protein Interaction, molecular interactions
- **Operations research data:** logistical networks, social networks, epistemological networks

System Data analysis

- **Network data in Computer science:** graphs are used to represent networks of communication, data organization, computational devices, the flow of computation, etc.
  - **Link analysis:** the link structure of a World Wide Web, Internet, Computer Network
  - website can be represented by a directed graph,
  - the vertices represent web pages and directed edges represent links from one page to another
- **Social network data analysis:**
  - graph with the structure of relationships between social entities.
  - entities are persons, groups, organizations, nation states, web sites, or scholarly publications

תודה רבה
Hebrew

Ευχαριστώ
Greek

Danke
German

Спасибо
Russian

धन्यवादः
Sanskrit

Merci
French

ধন্যবাদ
Bangla

நன்றி
Tamil

شكرا
Arabic

ಧನ್ಯವಾದಗಳು
Kannada

Thank You
English

നന്ദി
Malayalam

Grazie
Italian

多謝
Traditional Chinese

ధన్యవాదాలు
Telugu

આભાર
Gujarati

ਧੰਨਵਾਦ
Punjabi

धन्यवाद
Hindi & Marathi

Gracias
Spanish

多谢
Simplified Chinese

https://sites.google.com/site/animeshchaturvedi07

Obrigado
Portuguese

ありがとうございました
Japanese

ขอบคุณ
Thai

감사합니다
Korean