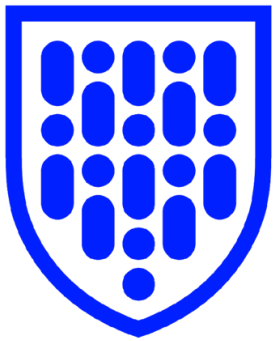


Institute of Data

2023

Data Science and AI

Module 9 Part 1

Artificial Intelligence (AI)

Agenda: Module 9 Part 1 - AI

- What is Artificial Intelligence (“AI”)?
- What are the different types of AI?
- History of AI
- Reinforcement Learning
- Multi-Agent Systems

What is Artificial Intelligence ("AI")?

- Homo sapiens means "**man the wise**" referring to our intelligence as the most defining features of us.
- "Human intelligence" refers to our ability to **perceive, understand, predict, and manipulate the world**.
- Intelligence is the ability to **acquire and apply knowledge and skills to achieve goals**.
- AI is usually defined in terms of **human's behaviour** or in terms of **mathematical and engineering** definition of "rationality".
    - Turing Test is an example of defining AI in terms of human behaviour
    - Theorem proving and building autonomous robots are examples of mathematical and engineering definitions.

What is Artificial Intelligence (“AI”)?

- AI is the study of *general principles of* ***rational agents*** *and on* ***components*** *for constructing them*.
- One of the most definitive text books on AI is: ***Artificial Intelligence A Modern Approach, by Stuart J. Russell and Peter Norvig*** details these principles and components.

What is Artificial Intelligence (“AI”)?

- AI systems demonstrate at least some of the behaviours associated with human intelligence.
  - **Planning**
  - **Learning**
  - **Reasoning**
  - **Problem-solving**
  - **Knowledge representation**
  - **Perception**
  - Motion
  - Manipulation
- To a lesser extent
  - Social intelligence
  - Creativity

AI – Key objectives

- Formulate **search** problems and implement search algorithms using admissible **heuristics**.
- Formulate **constraint satisfaction problems** and find solutions using constraint graphs.
- Describe **games** as **adversarial search problems** and implement optimal and efficient solutions.
- Formulate **nondeterministic** search in **reinforcement learning contexts**.
- Define the **machine learning** problem and implement simple algorithms including Naive Bayes, neural nets, and clustering.

What are the different types of AI?

- AI can be separated at a high-level into two broad types: **Narrow AI** and **General AI**
- Narrow AI is what is available in computers today: intelligent systems that were taught or learned how to execute specific tasks without being explicitly programmed to do so
  - Language and speech recognition on the Apple iPhone with Siri virtual assistant
  - Vision-recognition systems on self-driving cars
  - Product suggestion by recommendation engines on what one might like based on what previous purchases
- These systems are limited to learn or be taught to do **specific tasks**, hence the name Narrow AI

What are the different types of AI? Narrow AI

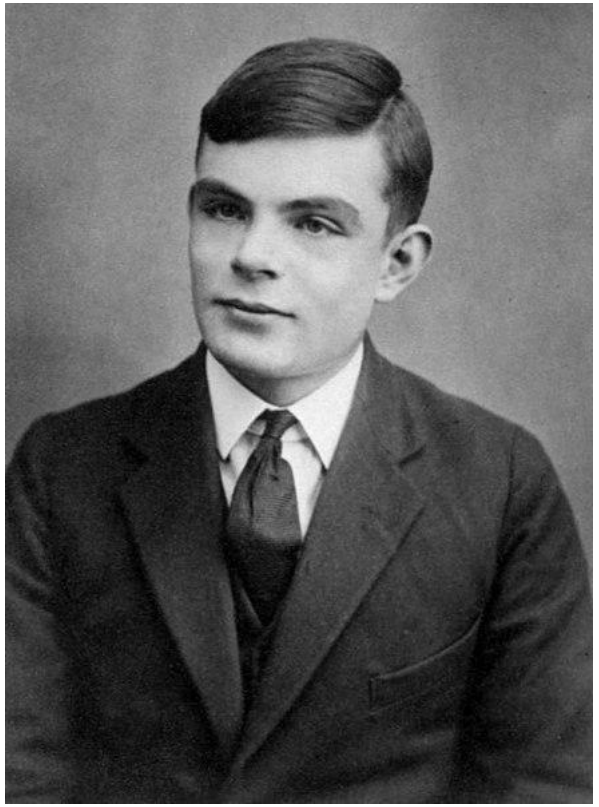
- A few other applications of Narrow AI
  - **Recommender engines**
  - **Chatbots**
    - Responding to **simple** help desk **queries**.
  - **Digital Assistants**
    - Alexa, Hey Google, Cortana, Siri
    - Organising personal and group **calendars**.
  - ML and **Cognitive** analytics assistants
    - Watson, Data Robot
  - Performing **visual inspections** of infrastructure by interpreting video input from drones.
  - Support the **identification** of potential tumours in **X-rays**.
  - Flagging inappropriate content online.
  - **Detecting wear and tear** in lifts from data gathered by IoT devices.

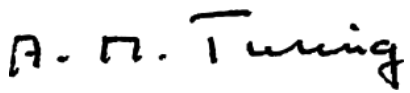
What are the different types of AI? General AI

- **Artificial General Intelligence (AGI)** is the similar to adaptable intellect characteristic of humans.
- AI as seen in movies, the likes of “HAL 9000” (2001, A Space Odyssey) or “Skynet” (The Terminator), which does not exist today.
- A form of intelligence capable of identifying and executing a **diverse set of tasks.**
  - Anything from controlling traffic to doing shopping, or to reason about a large roll of topics based on its previous experiences.
- Any projections are mostly **guesses** given the limited understanding of the human brain.

When did AI research start?

- After WWII, some people started to work on intelligent machines independently
- **Alan Turing**, an English mathematician, may have been the first
  - An early reference is a lecture given by Turing in **1947**
- Turing also may have been the first to suggest that the best approach to AI was via **software** instead of hardware
- There were many researchers on AI by the late 1950s, and the majority were programming computers in their work as its base
- A film about Alan Turing: The Imitation Game





Reference: Wikipedia

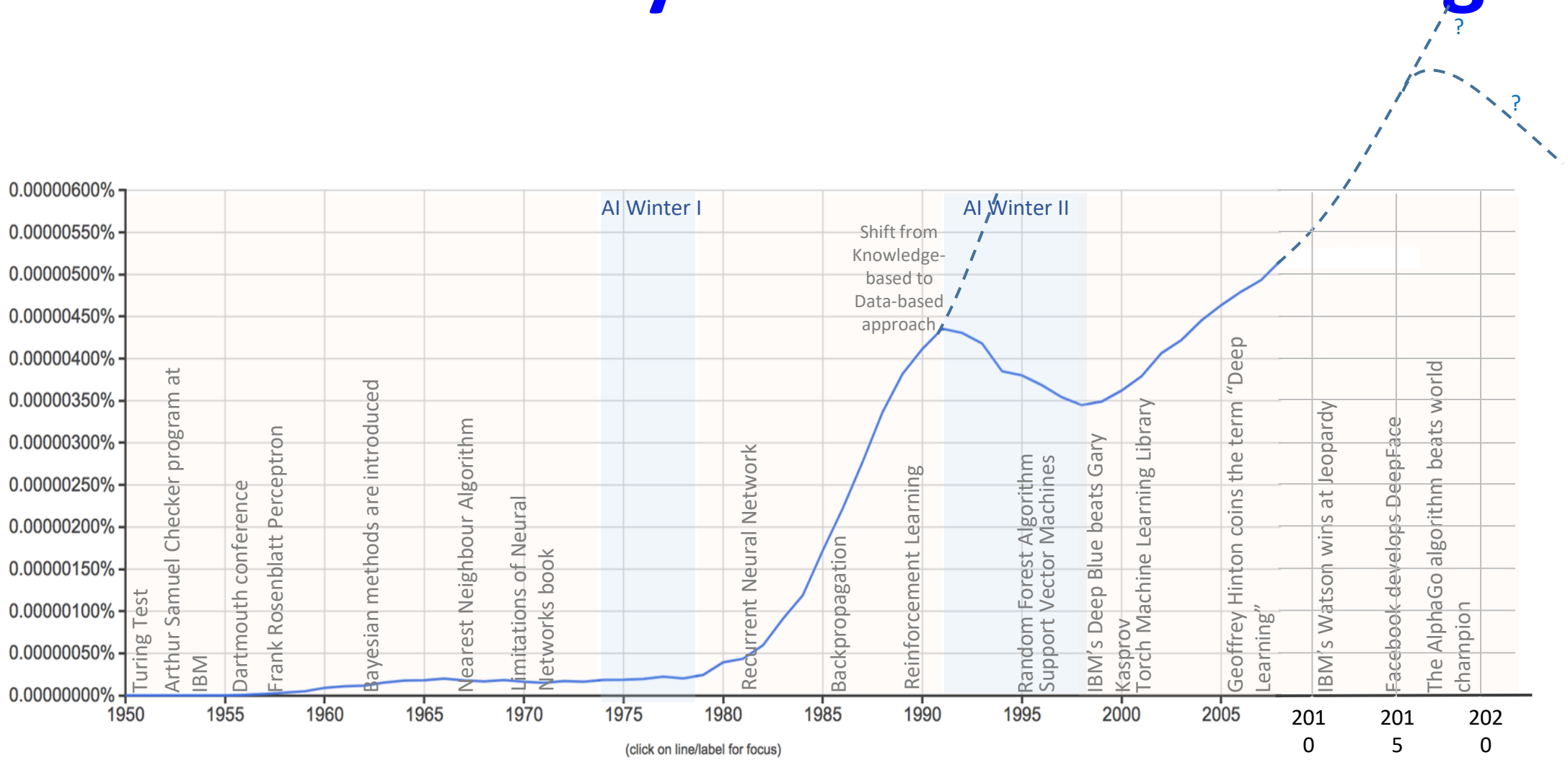
The Turing Test

- Alan Turing discussed the conditions to consider a machine to be intelligent in 1950
  - Alan Turing wrote a paper called **“Computing Machinery and Intelligence”** on the topic of artificial intelligence, published in 1950 in Mind Journal.
- If a machine could **imitate** a human with success to a knowledgeable observer, then you certainly should consider it intelligent.
- The Turing test is one-sided, as a machine that passes the test should certainly be considered intelligent without knowing enough about humans.
- This test would **satisfy** most people but **not all philosophers**.

AI Winter

- The Department of Defence heavily **funded** AI research in the U.S. in the middle of the **1960s.**
- AI's founders were **optimistic** about the future
  - "*Machines will be capable, within **twenty years**, of doing any work a man can do*"
  - "***Within a generation** ... the problem of creating 'artificial intelligence' will substantially be solved*"
- They **failed to recognise** the difficulty of some of the **remaining tasks**.
- In 1974 progress slowed and both British and the U.S. governments stopped exploratory research in AI.
- The next few years would later be known as "AI winter".

A data-driven story of Machine Learning



Google n-gram trend
Extrapolation

AI in the 1980s

- AI research revived due to the commercial success of **expert systems** in the early 1980s.
  - Expert systems are a kind of AI that simulated the **analytical skills and knowledge** of human experts.
- The market for AI reached over **a billion dollars** by 1985.
  - The British and U.S. governments restored funding for academic research inspired by **Japan's fifth generation** computer project
- AI fell into disrepute again after the collapse of the **Lisp Machine** market in **1987** and a went through a second hiatus.

AI in the 1990s

- AI began to be used for practical applications starting from the late 1990s and over the century in **data mining, logistics, medical diagnosis** and other areas.
- Most of the success is due to
  - Increasing **computational power**
  - Emphasis on **solving particular problems**
  - Experimentation on interaction between **AI and other fields** (like economics, statistics and mathematics).
  - Researchers embracing mathematical methods and scientific standards
- **IBM's Deep Blue** became the first chess-playing system to defeat a world chess champion on 11 May **1997**, chess master Garry Kasparov.

AI in 2000 and Onwards

- Algorithmic improvements, faster computers and access to vast amounts of data enabled advances in perception and **machine learning.**
- IBM's question answering system, named Watson, defeated in 2011 the two greatest **Jeopardy!** champions, Brad Rutter and Ken Jennings in an exhibition match.
- By 2012, data-hungry **deep learning** methods started to dominate accuracy benchmarks.
- **AlphaGo** become the first Go-playing system to beat the Go champion Lee Sedol on March 2016.

AI in 2000 and Onwards

- In 2017's Future of Go Summit, AlphaGo won a three-game match with Ke Jie, who was the world number 1 for two years
  - As Go is a game even more complex game than Chess, this achieves a significant milestone in the development of Artificial Intelligence
- Bloomberg's Jack Clark, said: **2015 was a landmark year** for artificial intelligence as
  - Error rates in image processing tasks have dropped massively since 2011
  - An increase in affordable neural networks, with the abundance in cloud computing infrastructure and an increase in datasets and research tools
- **China** greatly accelerated its funding by 2016 given its rapidly increasing research output and its vast supply of data
  - Some observers claim that China will become an "AI superpower"

Responsible AI and ethics

- Several frameworks for the ethics of AI have been proposed in recent years.
  - 2017: 23 Asilomar AI Principles
  - 2019: OECD AI Principles
  - 2021: UNESCO adopted a global standard
- The behaviour of both the artificial systems and the humans designing or using them need consideration.

Responsible AI and ethics - challenges

- lack of **transparency** – some algorithms are not explainable
- ethnic and gender **bias** – algorithm designers will have inherent assumptions and biases, training sets may lack diversity
- potential of unexpected behaviour due to poor generalisation performance
- unintended consequences of optimising a function
- impacts on our **interactions** and relationships with other humans
- devaluing people or replacing people in positions that require **empathy**
- **mass surveillance** and **privacy** issues
- AI **governance** and **accountability** - who takes responsibility for failure of an AI system?
- unreliable AI in **law enforcement**

Discussion

- What are some of the biggest **risks** you see with AI now and into the future?

Institute of Data

AI - intelligent systems

- Reinforcement Learning
- Multi-Agent Systems

Reinforcement Learning (RL)

- Reinforcement learning (RL) is an area of **Machine Learning**
- RL tries to take suitable action to maximise **reward** in a particular situation
- It is employed by various software and machines to find the best possible behaviour or path it should take in a specific situation
- Reinforcement learning differs from the supervised learning
  - In supervised learning **the training data** has the answers so the model is trained with the correct answers
  - In reinforcement learning, there is no answer but the **reinforcement agent** decides what to do to perform the given task
- RL is an area of machine learning concerned with how **intelligent agent** sought to take actions in an **environment** so as to **maximise** some notion of cumulative **reward**

Elements of Reinforcement Learning (RL)

- **Input:** an initial state from which the model will start
- **Output:** possible outcomes as there are variety of solution to a particular problem
- **Training:** The training is based upon the input, the model will return a state and the **environment** will decide to **reward** or **punish** the model based on its output
- The model keeps **continues to learn**
- The best solution is decided based on the **maximum reward**

Applications of Reinforcement Learning (RL)

- RL has been used to **play games** at levels at or higher than human abilities. AlphaGo and AlphaZero
- RL can be used in **robotics** for industrial automation
- RL can be used in typical **machine learning** and data processing
- RL can be used to create **training systems** that provide custom instruction and materials according to the requirement of students

Multi-Agent Systems

- A Multi-Agent System (MAS) is a computerised **self-organised** system having multiple interacting **intelligent agents**.
- MAS can handle problems that are difficult or impossible for a single agent or a monolithic system to solve.
- Intelligence can include **functional, procedural, methodical** approaches, algorithmic search or reinforcement learning.
- Multi-agent systems studies may deliver a proper approach in applications such as **disaster response, online trading and social structure modelling.**
- MAS usually apply **Reinforcement Learning** among other techniques.

Multi-Agent Systems - Characteristics

- Essential characteristics of agents in a multi-agent system
  - Autonomy: agents at least self-aware, autonomous and partially independent
  - Local views: agents do not have a full system view, or the system is too complex for an agent use such knowledge
  - Decentralisation: no agent is in control, there is no master or central agent

Multi-Agent Systems - Concept

- Multi-agent systems are composed of **agents** and their **environment.**
- Multi-agent systems research commonly refers to **software agents**
  - However, the agents could equally well be **robots, humans or human teams**
- Agents can be grouped into types from **simple to complex**
  - **Passive agents (no goals),** such as obstacle, key or object in any simple simulation
  - **Active agents (simple goals),** like birds in migration, or hunter-prey in the prey-predator model
  - **Cognitive agents** (complex computations)
- Agent environments can be divided into
  - **Discrete**
  - **Continuous**
  - **Virtual**

Multi-Agent Systems - Concept

- Agent environments can also be organised by properties like
  - Accessibility: if it is possible to collect full details about the environment
  - Dynamics: how many entities are currently influencing the environment
  - Discreteness: if the number of possible actions in the environment finite
  - Determinism: if an action causes a definite effect
  - Episodicity: whether agent actions in certain periods influence other periods
  - Dimensionality: whether spatial attributes are an important characteristics of the environment and the agent considers dimensions to make decisions

Multi-Agent Systems - Concept

- An appropriate **middleware** typically mediates the actions of the agents
- This component offers a first-layer design abstraction for multi-agent systems, providing means to control resource access and agent coordination

Multi-Agent Systems - Organisation and Direction

- Multi-agent systems can provide **self-organisation, self-direction**, other control paradigms and related complex behaviours even when the individual strategies of the agents are simple.
- Using an agreed **language**, limited to the system's **communication protocol**, an **exchange of knowledge** by the agents may lead to a universal improvement.
- **Agent Communication Language (ACL)** and **Knowledge Query Manipulation Language (KQML)** are languages examples.
- MAS tend to find the best solution without **intervention**.
- The systems also tend to **self-recovery**, be **fault tolerant** and restrict error propagation, mainly due to the redundancy of components.
- There is a high similarity to physical phenomena, such as energy minimising, where physical objects tend to achieve the lowest energy possible within the physically constrained world.

Discussion

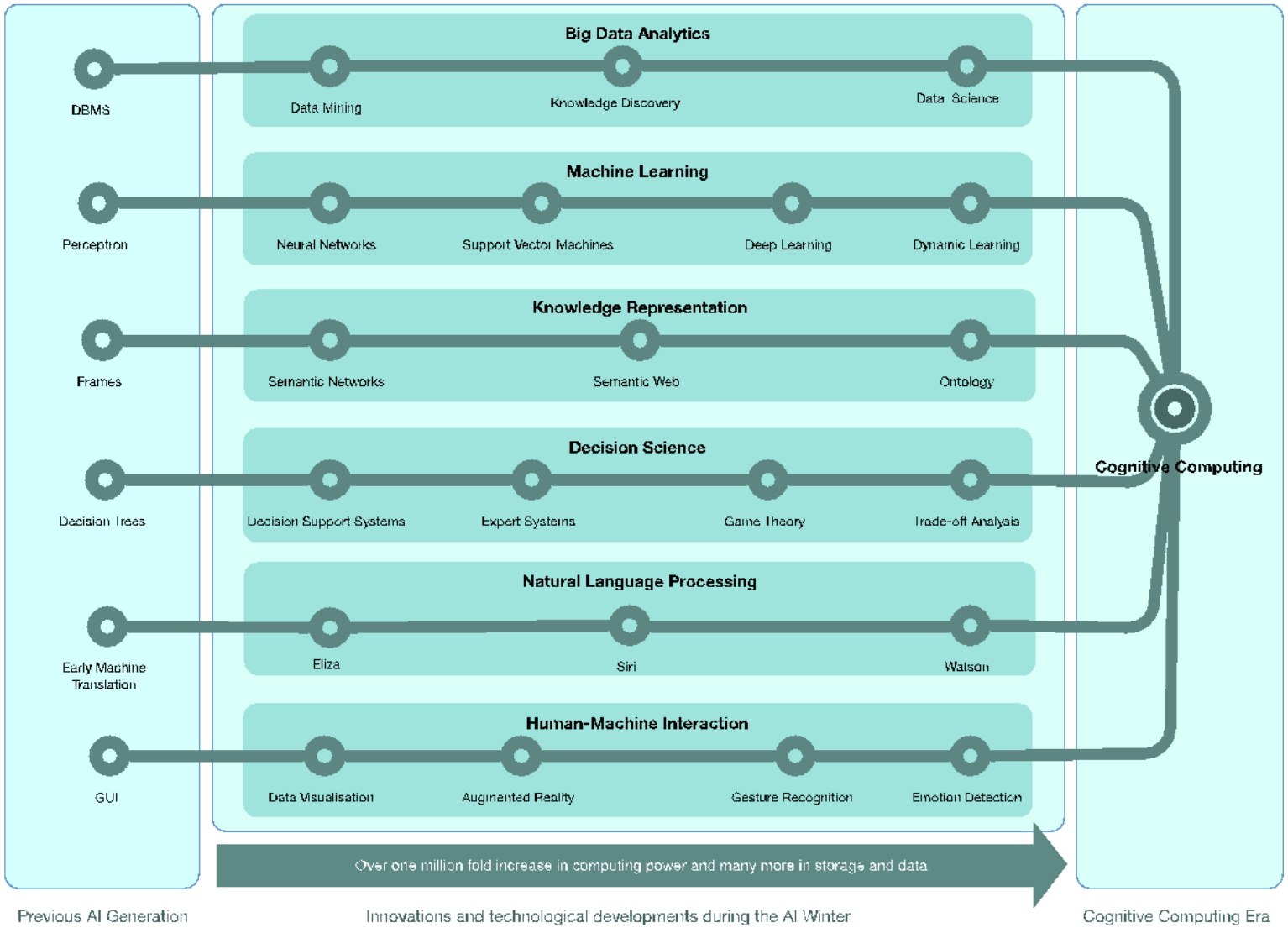
- What kinds of **applications** would you associate with AI now?
  - **Recommender engines**
  - **Chatbots**
  - **Digital Assistants**
    - Alexa, Hey Google, Cortana, Siri
  - ML and **Cognitive** analytics assistants
    - Watson, Data Robot

- What questions or comments do you have about the future of AI?

Questions

Appendices

Cognitive Computing



End of Presentation!