Headquarters U.S. Space Force

Everything You Ever Wanted to Know About Data in 50 Minutes

Ground Systems Architectures Workshop

February 28, 2022

Mark Brady

HQSF Chief Data Officer

Overall Classification:
UNCLASSIFIED

Past and Present Historical Trends in US Warfare

- Last peer to peer WWII ends 1945
- 76 years of asymmetrical warfare in the interim
  - Korea
  - Vietnam
  - Iraq
  - Afghanistan
  - Etc.
- Prospects of peer to peer now loom again
  - Europe
  - Taiwan
- Greatest challenge in 76 years
- How to win?

It Takes a Winning Strategy: The Strategist

It Takes a Winning Strategy: Five Factors Completely Determine the Strategic Universe

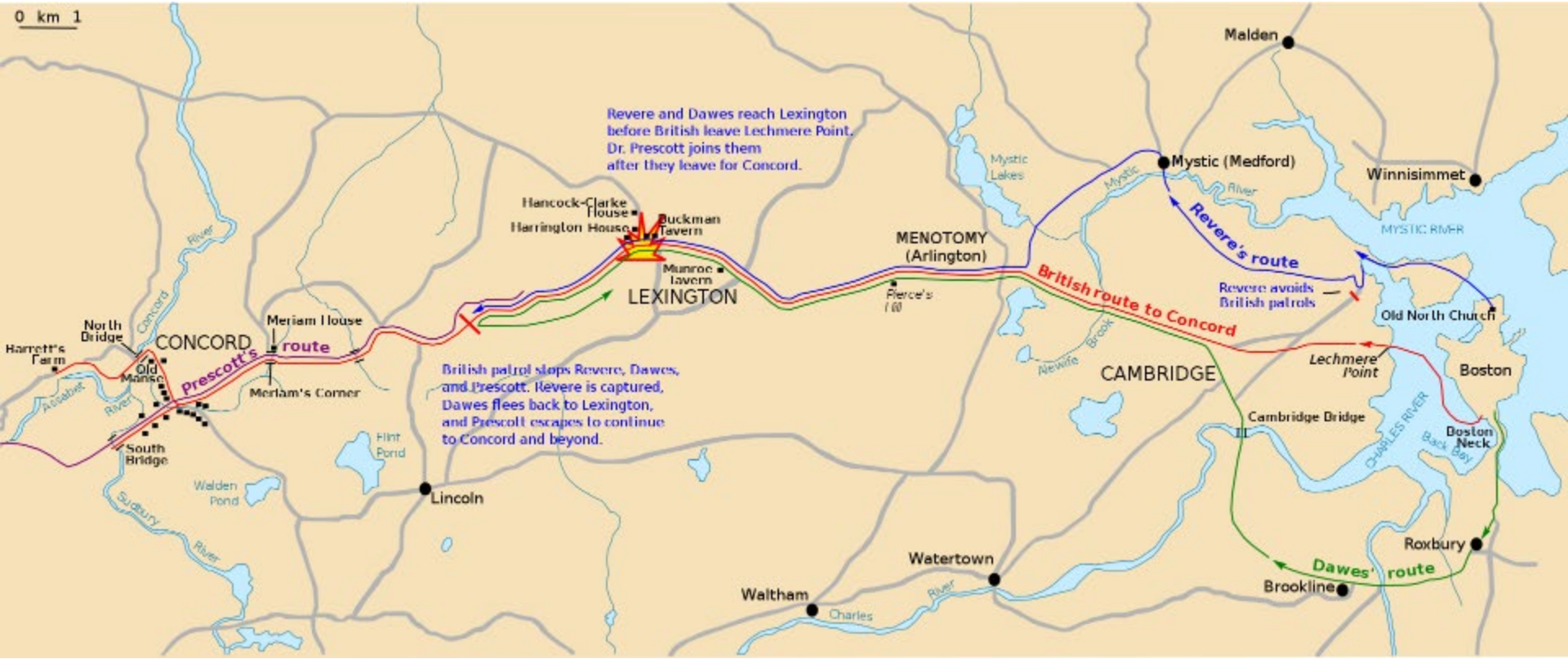
1. What you have
2. What you know
3. What the opposition has
4. What the opposition knows
5. The environment

What You Know as a Lens on the Other Four factors

- Do you know what you have?
- Do you know what the opposition has?
- Do you know what the opposition knows?
- Do you know the environment?

- **There is even, do you know what you know?**
- **Turns out to be profound.**

Historical Illustration of the Five factors
Lexington & Concord



The Players – All British

- Red Coats
- Patriots
- Loyalists

Five Factor Analysis Lexington - Concord

- Patriots have – Weapons and powder cache at Concord
- Red Coats have – 700 regulars
- Red Coats know – There is a cache at Concord
- Patriots know – 700 regulars are coming

- Five factor interactions
  - Red Coats knew what the Patriots had
  - Patriots knew what the Red Coats knew

- Red Coat strategy becomes a false strategy when what they knew turned from true to false. Recall the lens.

Historical Examples (Continued)

Common Misconception: Enough knowing & having will guarantee success

Do you know what you know?

Four Levels of Knowing

1. Having data.
2. Knowing what data you have.
3. Understanding your data by itself.
4. Understanding your data in the context of other data.

Must be at Level 4 for a peer-to-peer

Getting to Level 4 Knowing: User Models

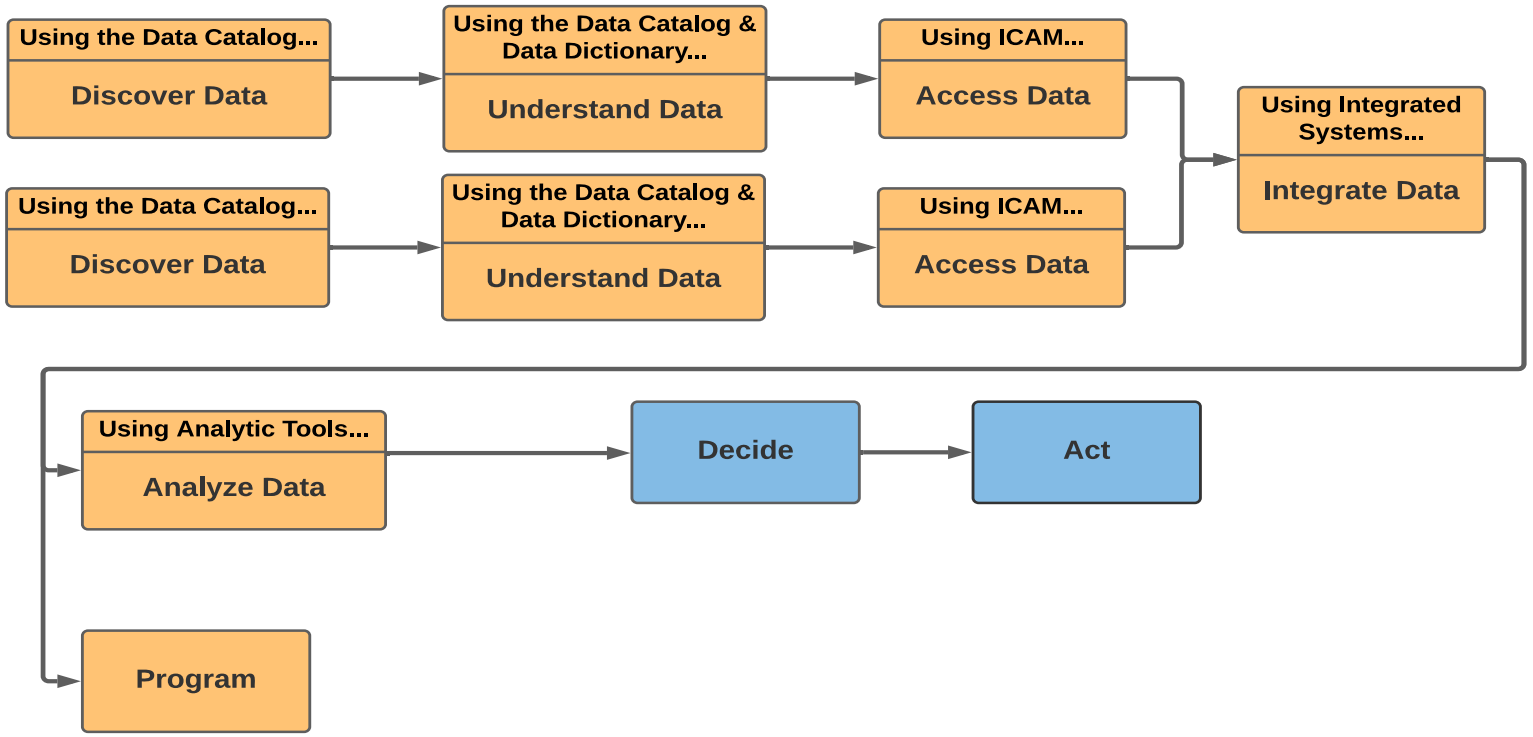
The underlying foundation of next generation data management are the data user models.

These user models inform us regarding activities data users must carry out in support of an organization's mission:

- The Active User Model
- The Notified User Model
- The Command and Control User Model

The Active User Model

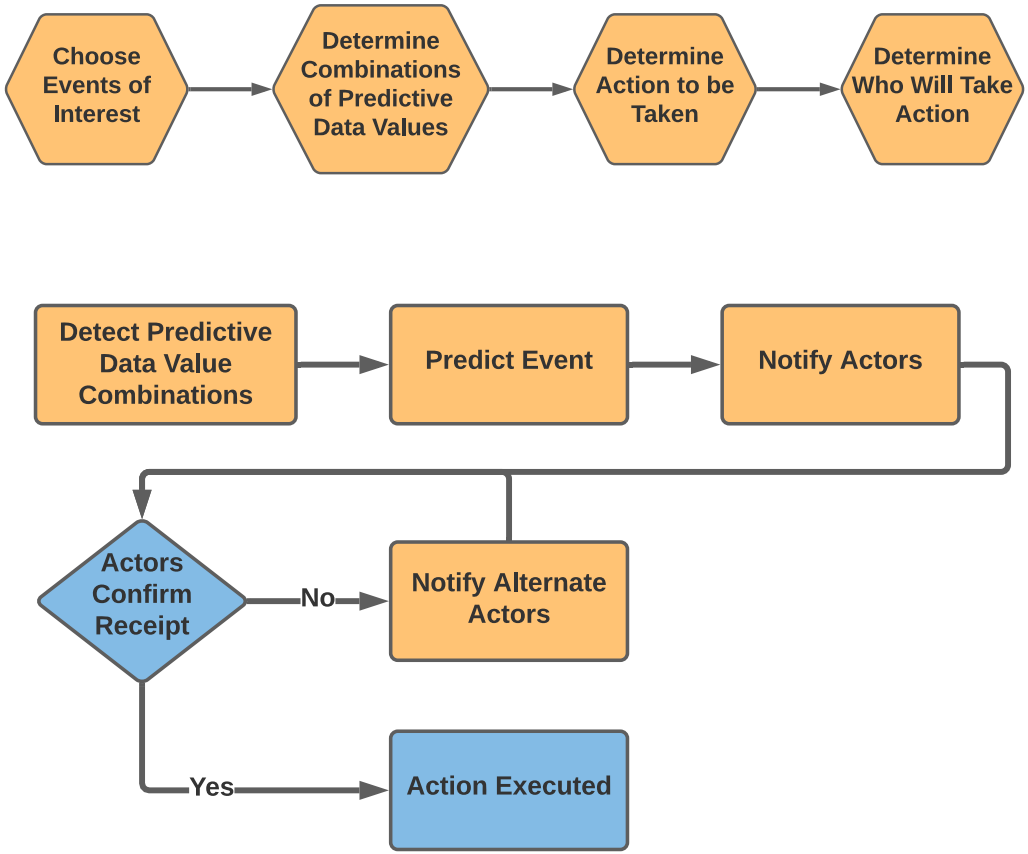
Helping a user discover and understand dozens, hundreds, or thousands of their organization's data assets.



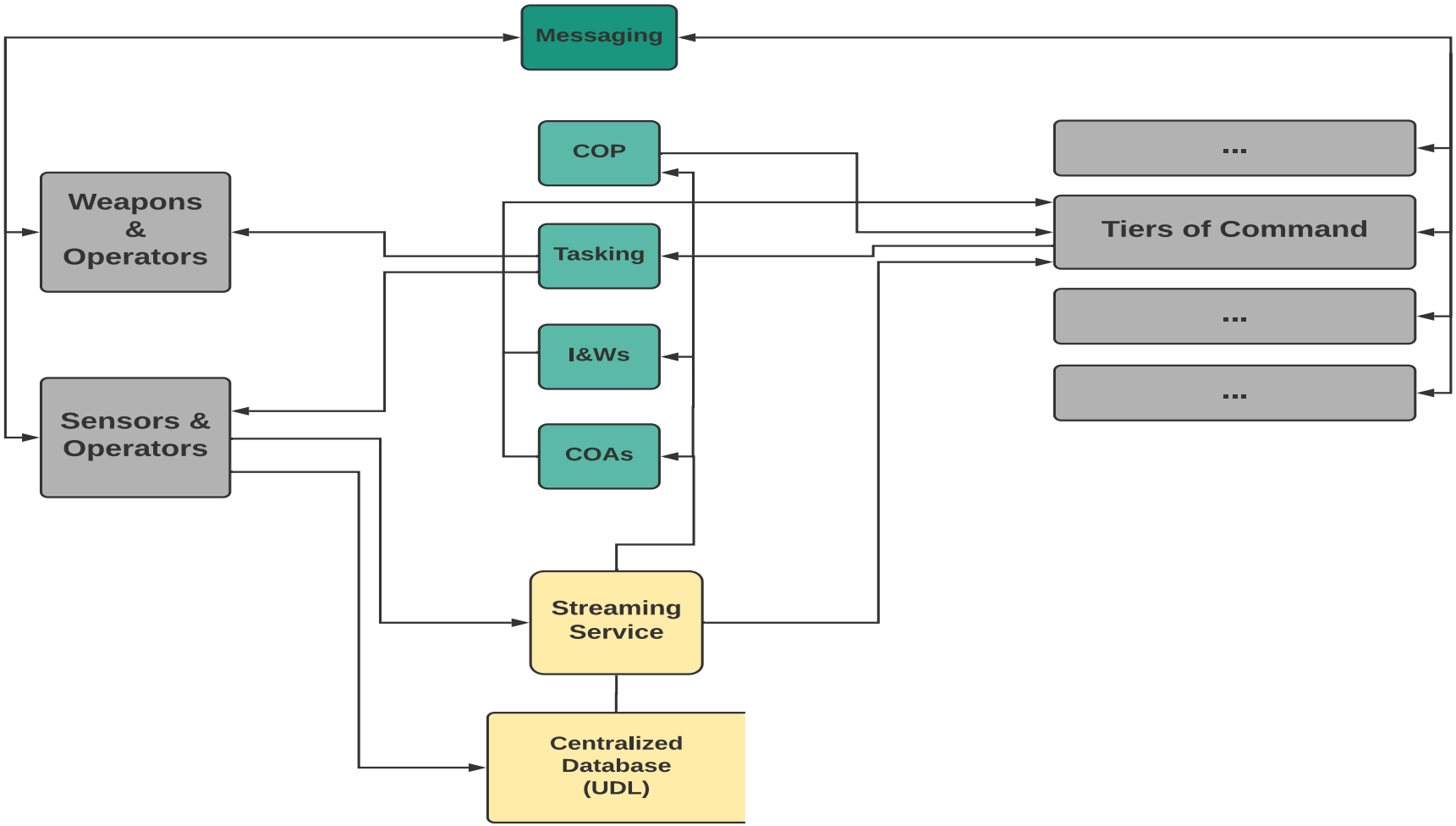
The Notified User Model

Unlike the active data user, the notified user is not actively seeking insights from the organization's data assets.

Rather, the data assets are sent to user via a data system called a notification system to inform advanced planning.



The Command and Control User Model

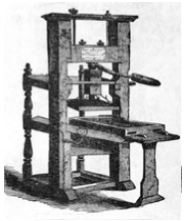


Communicating About Data: Bridging the Chasm

The Post Literate Era

Earliest Written Communication

Invention of the Printing Press



40% of Americans Read Literature

3500 BC | 23 BC | 1430s | 1920s | 1940s | 2020s

First Books

Literacy Reaches 70% in Parts of US

Data Steward Fails to Document Database

- Avoidance of writing and reading
- High levels of ambiguity

Terminology:
More Important Than You Thought

Data Life Cycle	Indexing	Transactional Data
Database	Metadata	Analytical Data
Data System	Data Fabric	Key
Ecosystem	Schema	Query
Dataflow	AI	Data Lake
Relational Database	Attribute	Normal Form
Column Store	Data Object	Join

Terminology: More Important Than You Thought

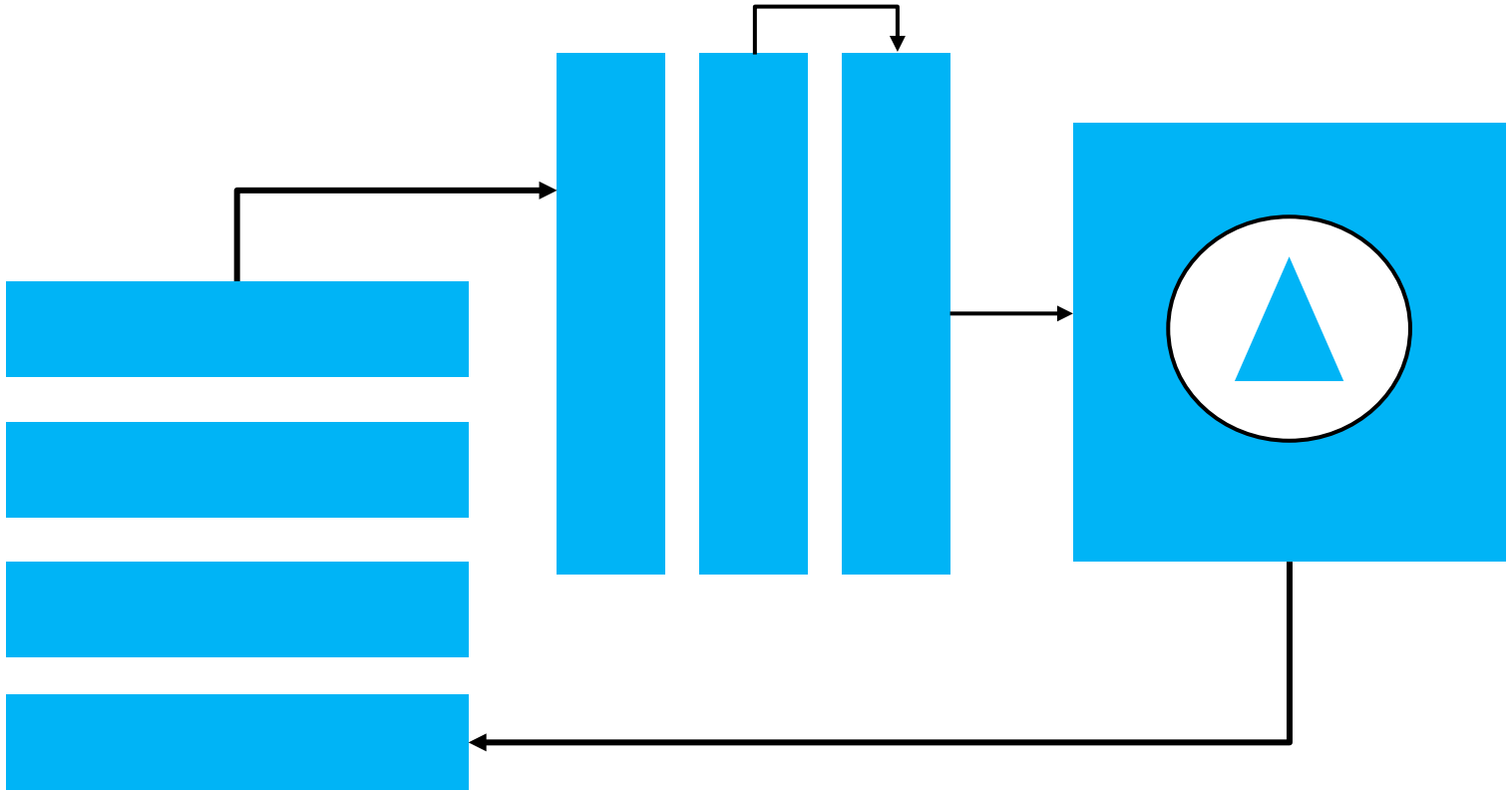
Data Life Cycle	Indexing	Transactional Data
Database	Metadata	Analytical Data
Data System	Data Fabric	Key
Ecosystem	Schema	Query
Dataflow	AI	Data Lake
Relational Database	Attribute	Normal Form
Column Store	Data Object	Join

- Undefined
- Potentially Problematic
- Inaccurate

PowerPoint as an Engineering or Proposal Document

- Complete sentences vs ambiguity

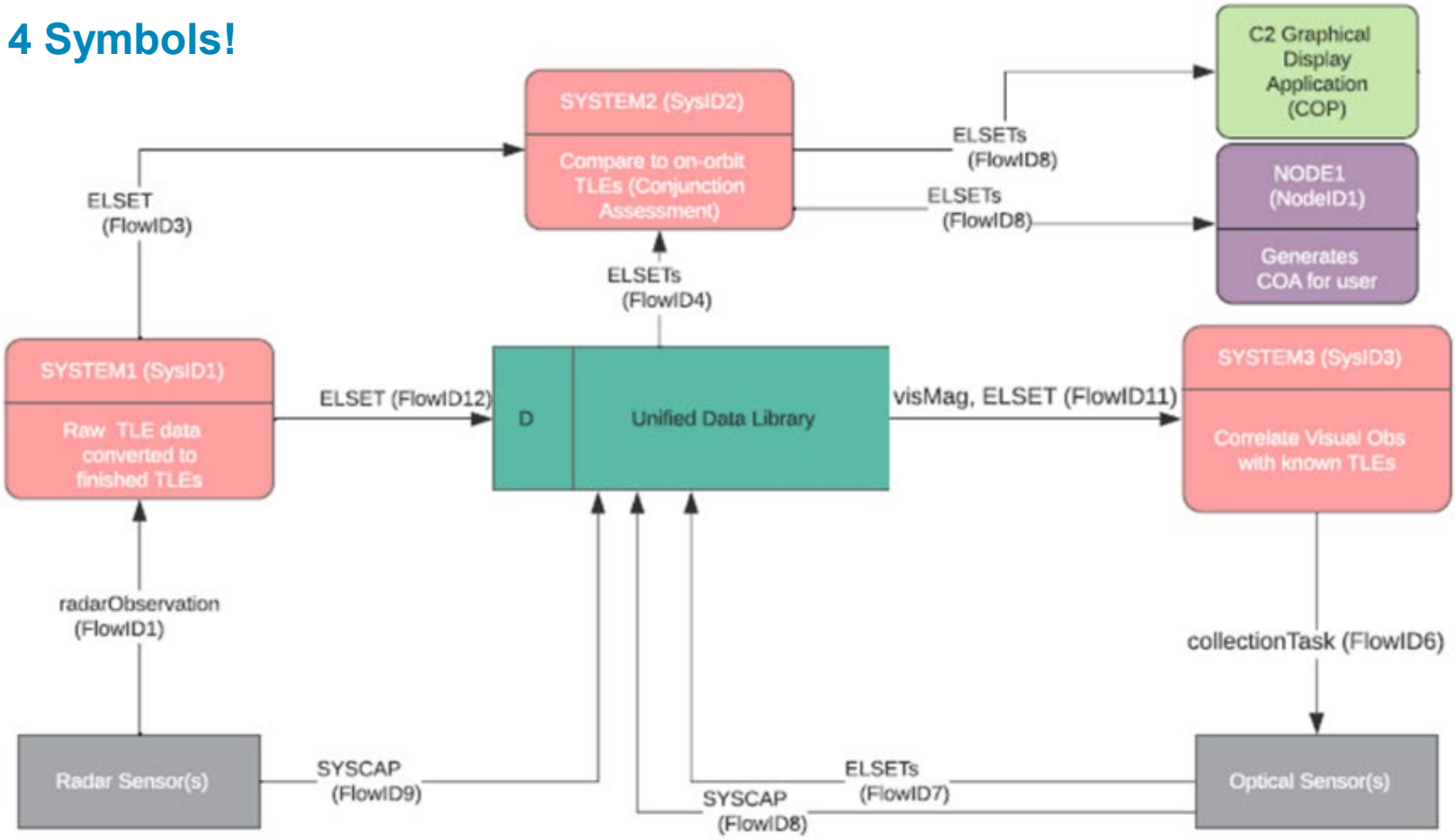
The Infographic as an Engineering or Proposal Document



Is a picture really worth a thousand words?

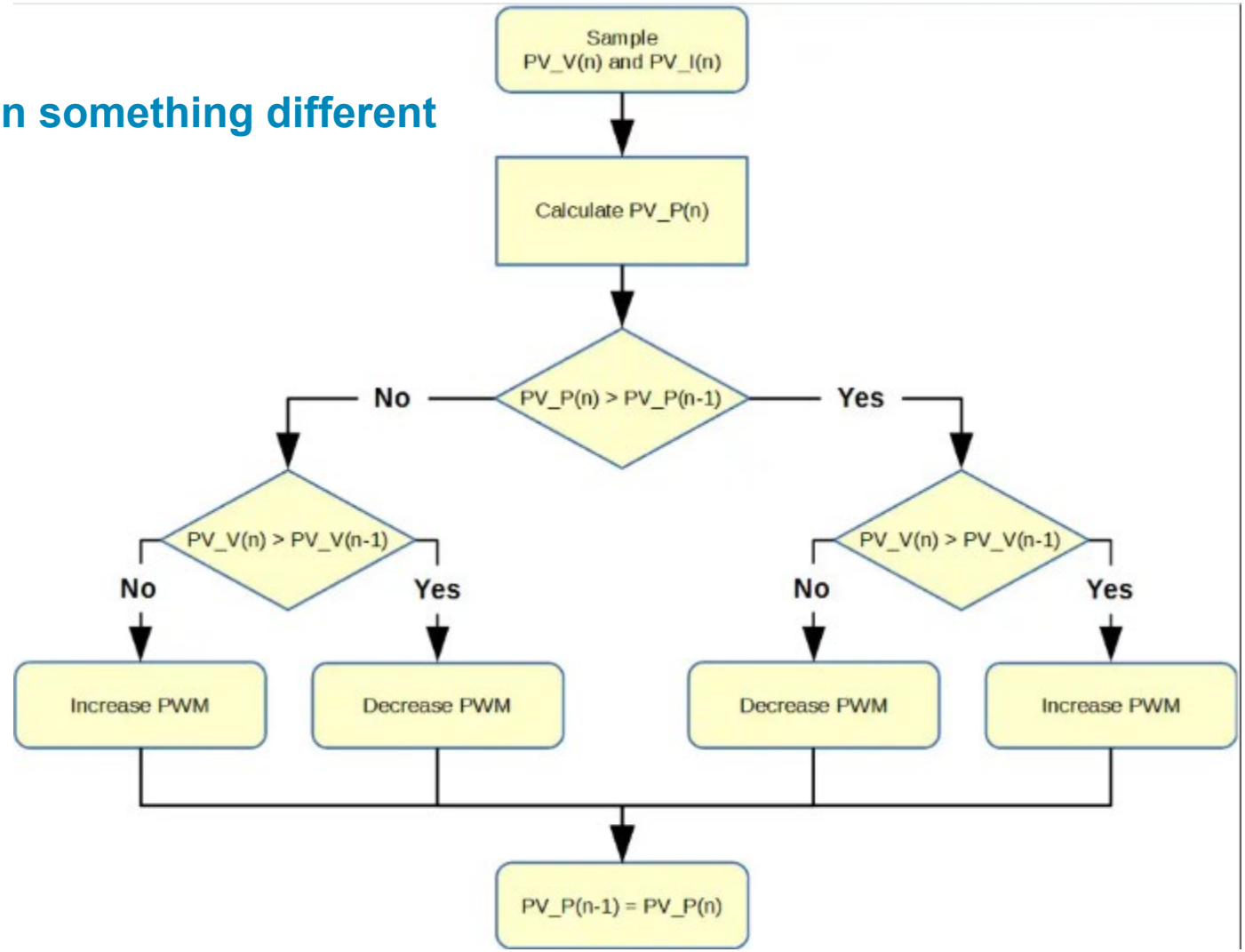
Unambiguous Diagrammatic Languages: The Dataflow

4 Symbols!



Unambiguous Diagrammatic Languages: Flowchart

5 Symbols!
Arrows mean something different



Oher Diagrammatic Languages

- Unified Modeling Language (UML)
- System Modeling Language (SysML)

- DODAF
  - Is not a diagrammatic language!
  - No syntax

How to Communicate Vital Technical Information to Anyone

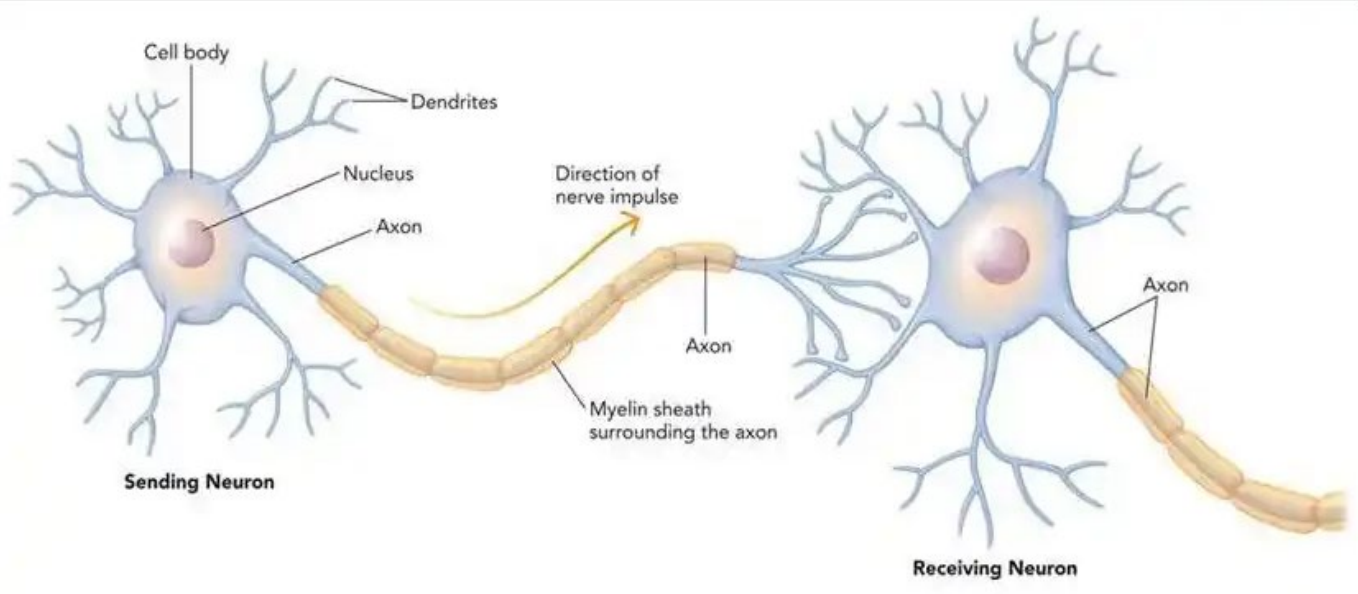
- Speaking
  - Learn and use well defined terminology
- Writing
  - Don’t be afraid to make your point
  - No more, no less
  - What did your reader learn?
- Don’t be afraid to read
- Diagrammatic Languages
  - Use them
  - Avoid infographics
  - Don’t mix
  - Minimize symbology

Artificial Intelligence

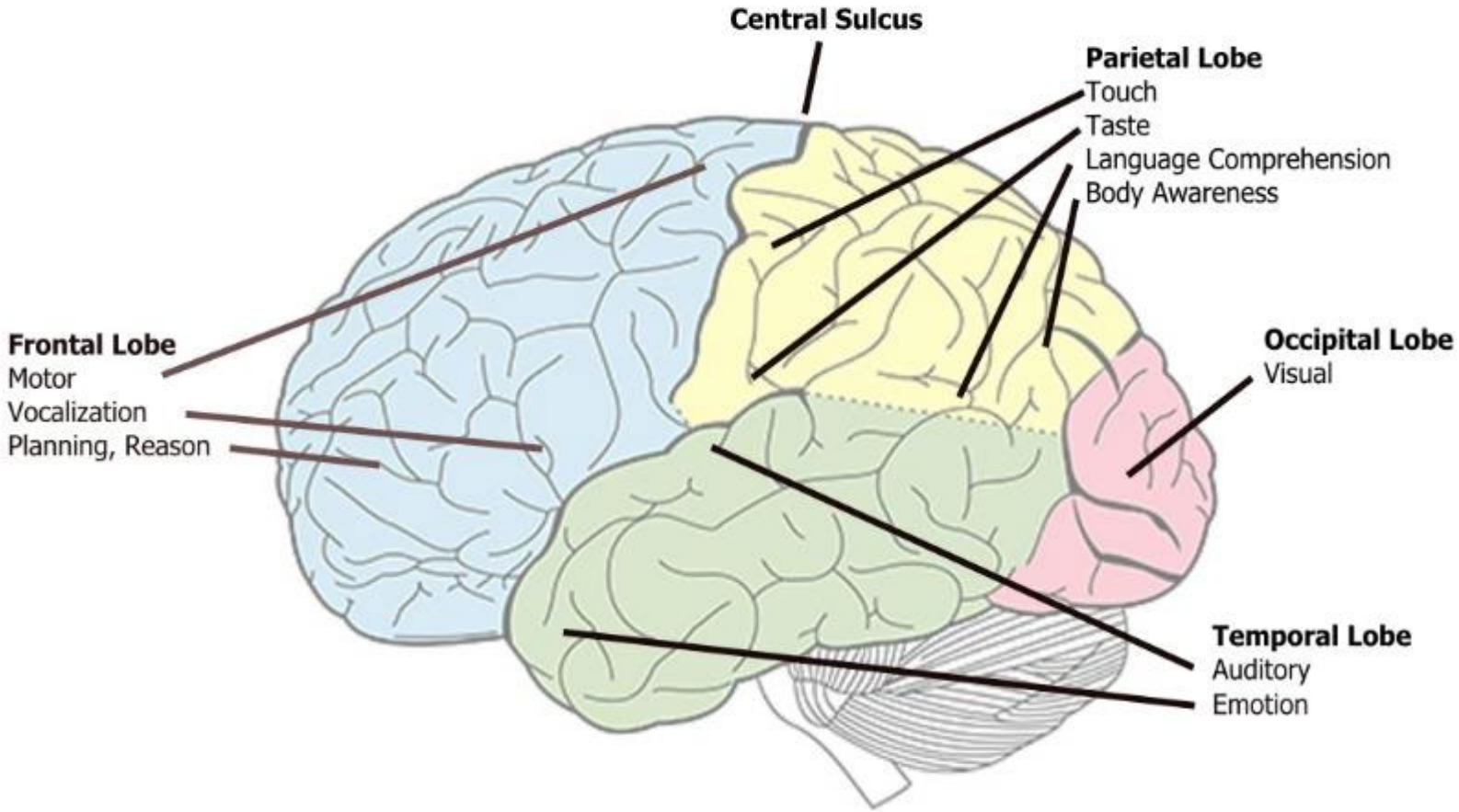
- Categories of intelligence
- The future of artificial intelligence
- How to use artificial intelligence

What is Intelligence?

- Some impressive data processing capability
- What's the threshold for impressive?
- If unsure, include all
- The biological model
- The pendulum swings from biomimetic to engineering approaches
- Neuronal level



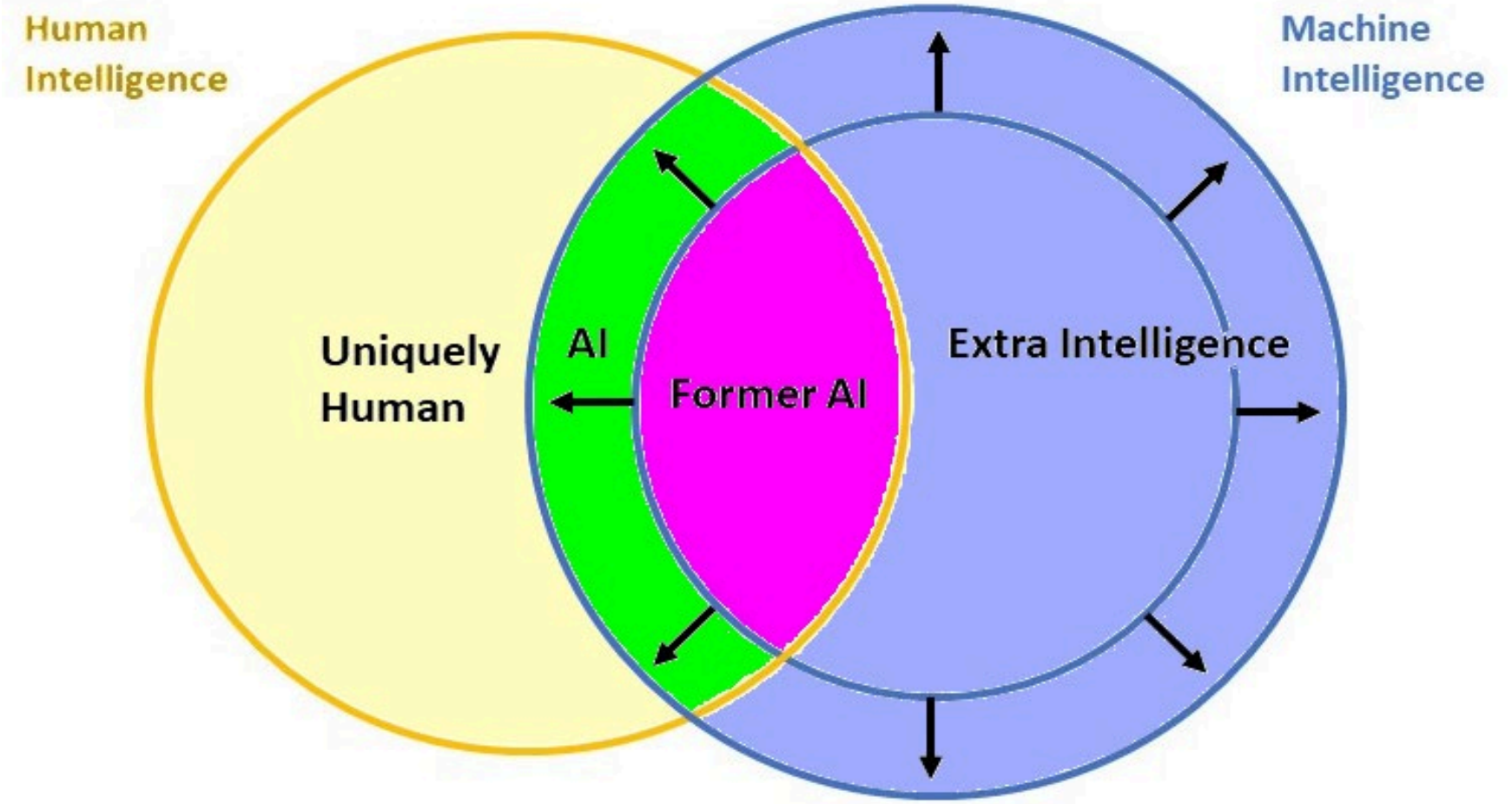
Biomimetic Variation: Functional Neuroanatomy



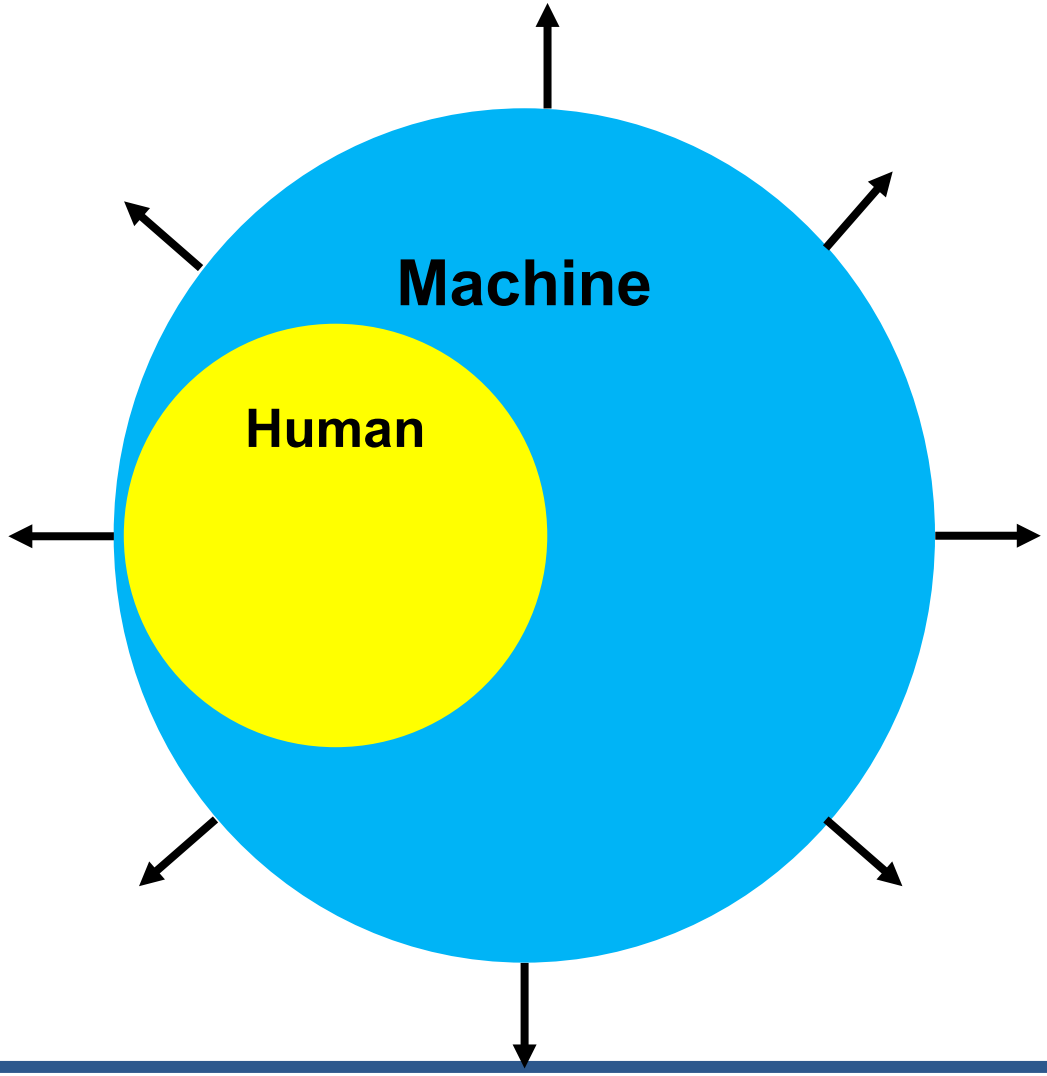
Functional Neuroanatomy Inspires AI Reseach

Human Intelligence	Corresponding AI Research
Sensation and Perception	• Computer vision • Speech recognition • Natural language processing
Motor	• Robotics • Speech generation
Planning and Reason	• Game theory • Autonomous vehicles
Emotion	• Objective function in machine learning

Intelligence Categories



The Singularity



How Will We get There?
Human vs Machine Comparison

Capacity	Human	Machine
Memory	• Associative • Fallible	• Concrete • Non-fallible
Calculations	• Error prone • Slow	• Perfect • Fast
Pre-programmed logic	• Error prone • Slow	• Perfect • Fast
Theorem proving	• Strong (for some)	• Weak
Analogical reasoning	• Strong (relates to associative memory)	• Weak
Real world experience	• Great	• Limited
Intuition	• High (relates to real world experience)	• None
Tolerance for variance	• High	• Low but improving
Autonomy	• High	• Low

Path to the Singularity

- Autonomy depends on **analogical reasoning** and intuition.
- Analogical reasoning depends on **associative memory**.
- Intuition depends on real world experience.
- Real world experience depends on **robotics.**

Dangers of the Singularity

- SkyNet
  - Elon Musk
  - Steven Hawking

What to Do?

- Asimov's 3 laws of Robotics
  - **First Law** - A robot may not injure a human being or, through inaction, allow a human being to come to harm.
  - **Second Law** - A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
  - **Third Law** - A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.
- The objective function, machine emotion
- There will be dangerous AI due to
  - Human malice
  - Evolving objective functions

- Create defensive AI to counter danger AI using objective functions
  - Protective of humans
  - Non-evolving

How to Use Current AI

- AI often relies on statistical methods and learning from data.
- That data may vary.
- For these reasons, results may have unpredictable flaws, somewhat like humans.
- Thus, don’t use AI where traditional predictable algorithms will work.
- Only apply where the algorithm is unknown to programmers.
- Apply when the data has structure not completely understood by humans.
  - Natural language!
  - Sensor data
- Don’t apply sloppy data management practices and then use AI to clean up the mess.

Q&A