Ai and AI

Artificial Intelligence and Automatic Indexing

A problem of Semantics and Definitions

DCMI 2021

- Marjorie Hlava

- Access Innovations, Inc.

- mhlava@accessinn.com

Access Innovations, Inc.

We Have a Definition Problem

- Taxonomy, Thesaurus, Ontology
- Incorrect use of these terms
- People use them interchangeably
- They are certainly NOT the same
- Artificial Intelligence and Machine Learning have a similar problem
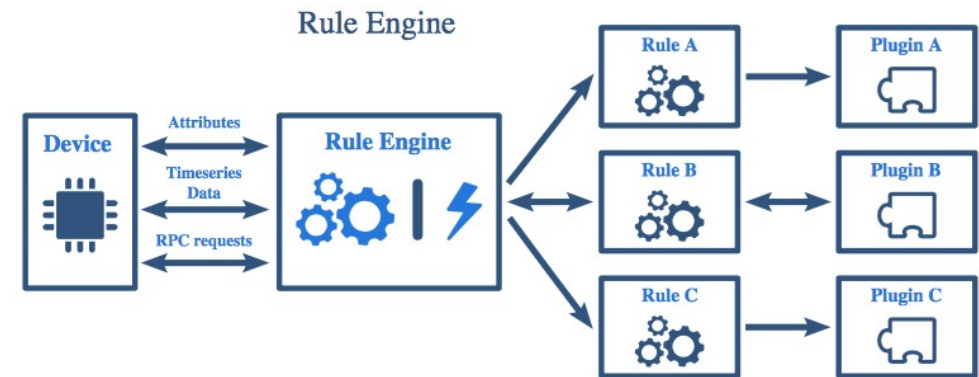
What is Automated Indexing or (Tagging)?

- The computerized process of scanning large volumes of documents against a controlled vocabulary, taxonomy, thesaurus, or ontology, and using those controlled terms to quickly and effectively index (tag) large electronic document depositories.

- https://en.wikipedia.org/wiki/Automatic_indexing

- Auto Cat (Automatic Categorization)

- Automatic Tagging

- Automatic Classification
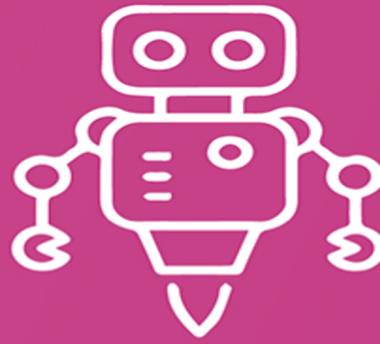
What is Artificial Intelligence?

- Artificial Intelligence is often defined as "The simulation of human intelligence process by machines, especially computer systems."

- Can be statistical, rules engines, or other forms of intelligence
  - Covers several domains
    - Natural language processing
    - Computer vision
    - Speech recognition
    - Object recognition

Rule Engine

Device — Attributes / Timeseries Data / RPC requests — Rule Engine — Rule A → Plugin A, Rule B ↔ Plugin B, Rule C → Plugin C

ARTIFICIAL INTELLIGENCE
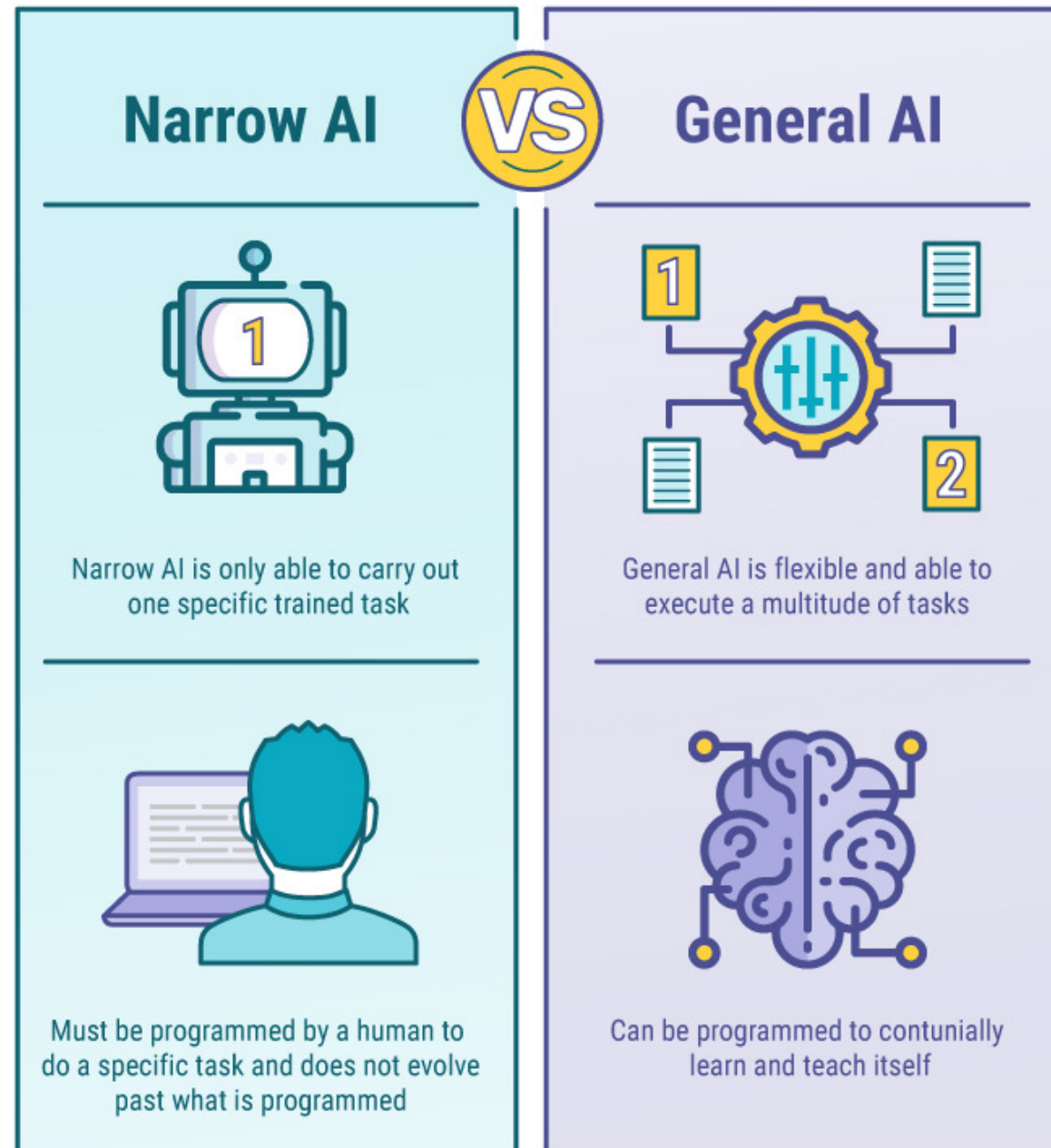
MACHINE LEARNING

DEEP LEARNING

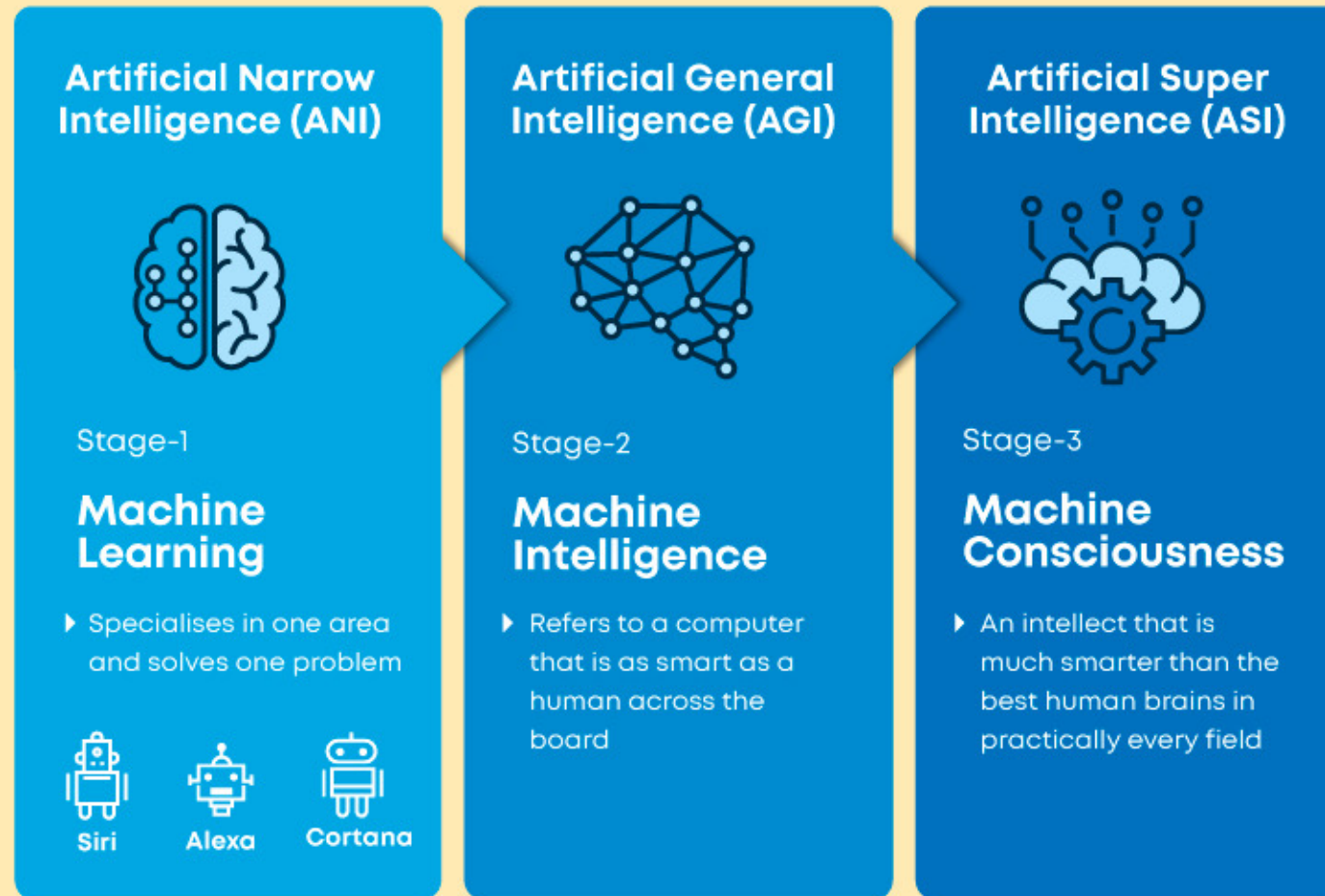# Artificial Intelligence vs Machine Learning

- Terms often used interchangeably
- One is often defined using the other
- They have overlapping domains and applications
- Machine learning is a subset of artificial intelligence
- Not all artificial intelligence is machine learning
  - Both tools often apply to the same domains, including natural language processing and computer vision
  - The differences are in how they are created and how they are applied to workflows requiring some form of decision making
  - With varied applications one may be more appropriate

2 Kinds of AI

Narrow AI VS **General AI**

Narrow AI is only able to carry out one specific trained task

General AI is flexible and able to execute a multitude of tasks

Must be programmed by a human to do a specific task and does not evolve past what is programmed

Can be programmed to contunially learn and teach itself

Access Innovations, Inc.

https://smartcompliance.co/blog/artificial-intelligence-improve-compliance

3 Types of Artificial Intelligence

Access Innovations, Inc.

3 Kinds of AI

Artificial Narrow Intelligence (ANI)	Artificial General Intelligence (AGI)	Artificial Super Intelligence (ASI)
Stage-1	Stage-2	Stage-3
Machine Learning	**Machine Intelligence**	**Machine Consciousness**
▸ Specialises in one area and solves one problem	▸ Refers to a computer that is as smart as a human across the board	▸ An intellect that is much smarter than the best human brains in practically every field
Siri Alexa Cortana		

7 Kinds of AI

10 Kinds of Functional AI

Quality Inspector
Machine Vision software

Personal Assistant
Siri, Google Now, Cortana

Clinician
DeepMind Health

Driver
Self-driving cars

Translator
Google Translate

AI is a

From Narrow to General AI
and From External to Internal Intelligence Peter Voss Oct 3, 2017

Writer
Narrative Sciences – Quill

Machine Operator
Baxter, Universal Robots

Stock Broker
Neokami

Soldier
Boston Dynamics - Petman

Chef
Moley Robotics

https://medium.com/intuitionmachine/from-narrow-to-general-ai-e21b568155b9

Access Innovations, Inc.

Hal, the computer in 2001: A Space Odyssey – was that Machine Learning or AI?

Access Innovations, Inc.
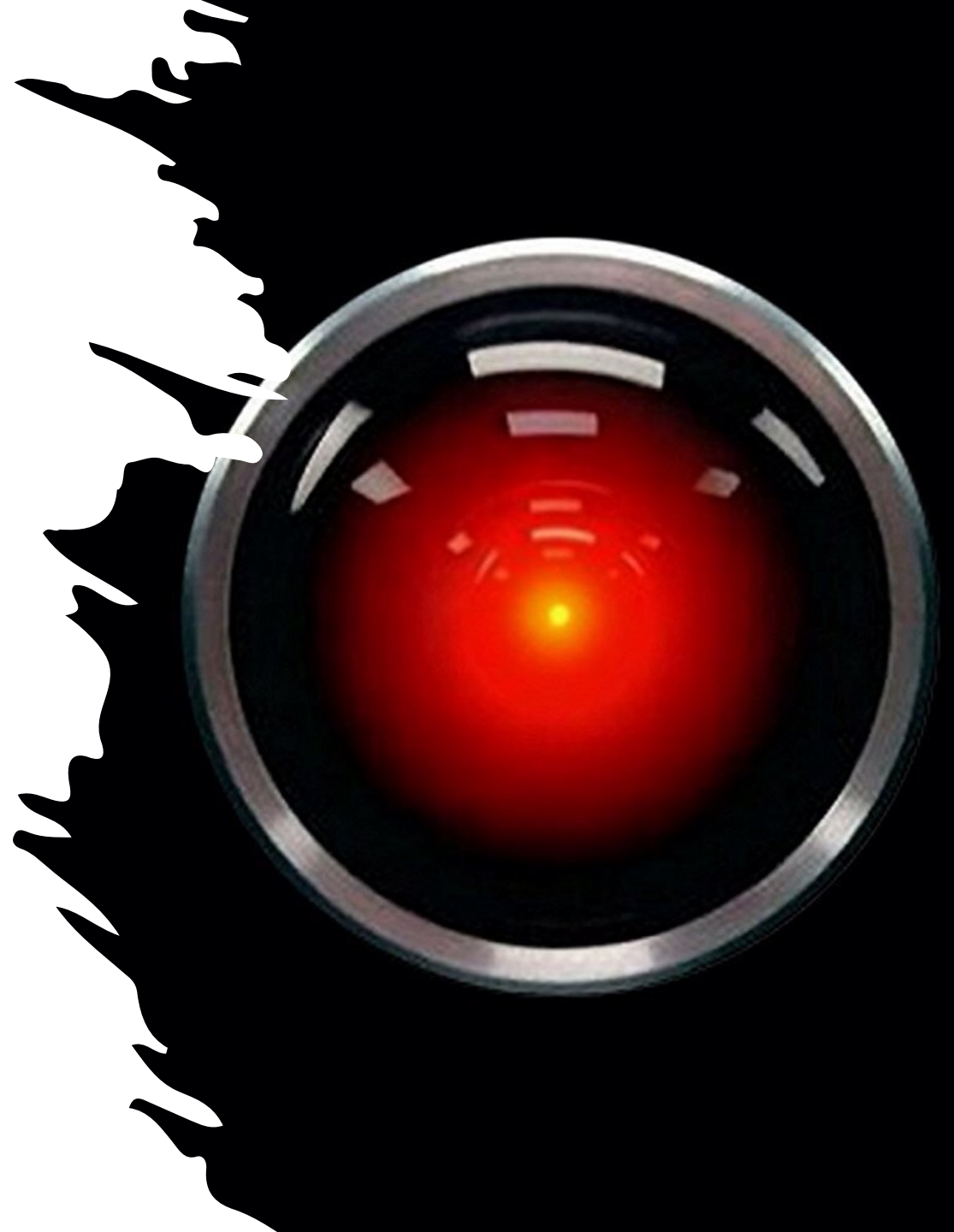
# Why the Fuss Over Definitions?

- **Artificial Intelligence,** in general, is very comprehensible for the developers to understand why it is working

- **Rules engines** use if/then and other advanced statements allowing the developer to very exactly define the system's behavior

    - If(cost < 1000.00) then purchase

Access Innovations, Inc.
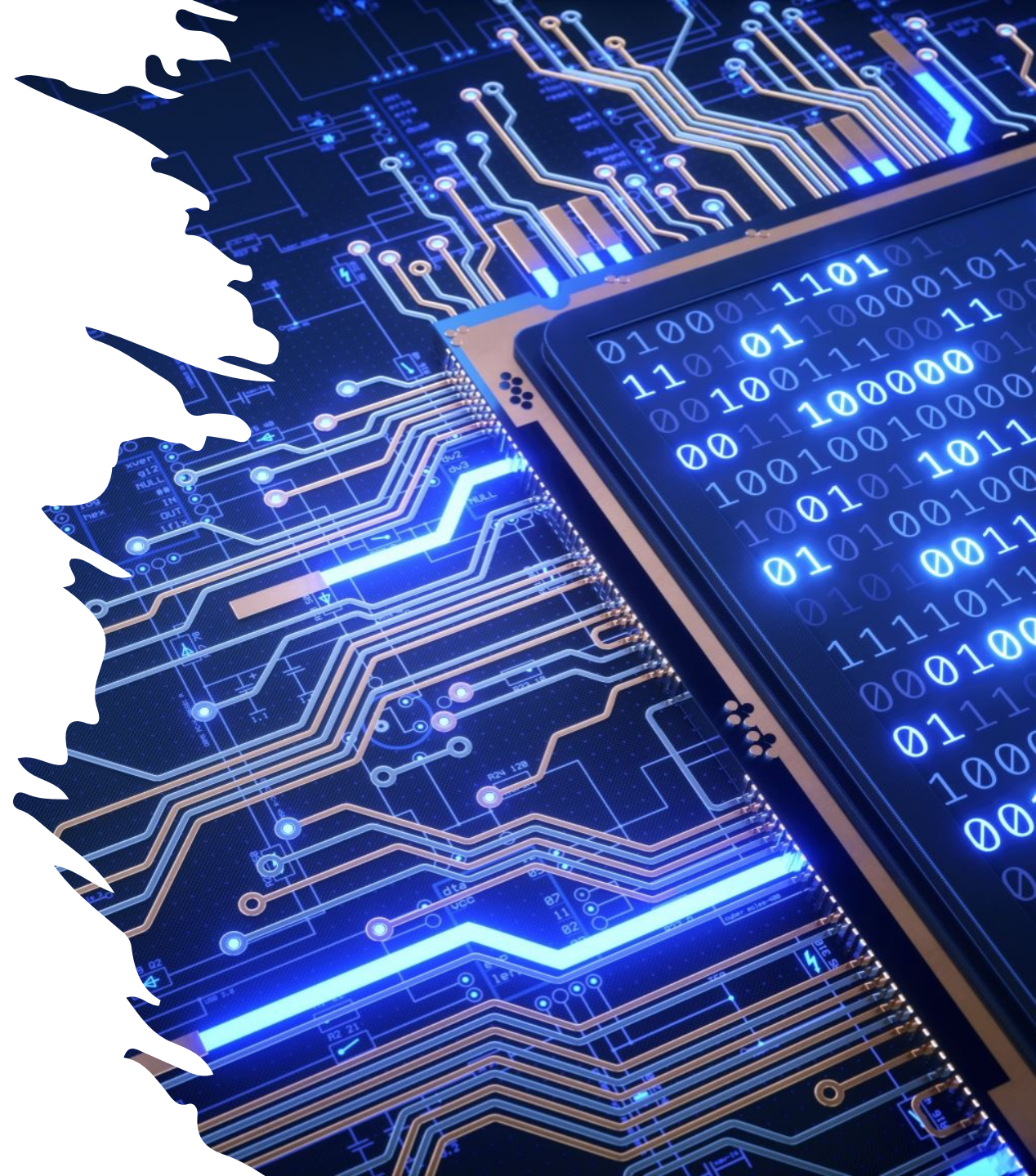
What is Machine Learning?
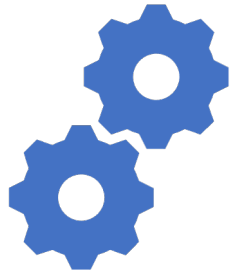
Machine Mearning is often defined as, "a method of data analysis that automates analytical model building."
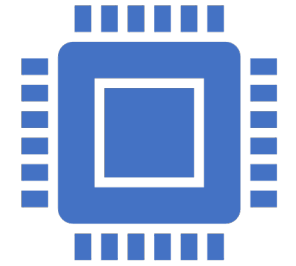
Purely derived by statistical (stochastic) means

Covers several domains as well

- Natural language processing
- Computer vision
- Speech recognition
- Object recognition
- Search engines

Machine Learning Models

- Collections of "neurons" that learn, through stochastic processes, how to accurately and precisely classify an input
  - For example, being able to repeatedly extract segmentations from a particular image
  - Very powerful
  - None can truly understand how the various neurons are coming to the conclusions that they are
  - This leaves very little room for the developer to be able to tweak how the system is behaving, aside from rough estimates around the hyperparameters of the system
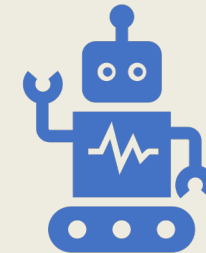
With These Distinct Differences In Mind...

Machine Learning limits clear understanding of how the systems are working

Artificial Intelligence allows replicable, repeatable and understandable applications

Use cases will dictate the application chosen: machine learning or a more malleable form of artificial intelligence

We have:

Entities (people, places, things)
> Identification
> Salience via weighting

Syntactic analysis (Parsing)

Semantic Analysis

Sentiment Analysis

Pragmatic analysis

Grammar

Lemmatization - stemming

Morphological

Lexical variations - synonyms

Part-of-speech tagging

Sentence boundary

Punctuation mark,

Abbreviations

Terminology extraction

Term weighting

Co-Occurrence*

Rules to increase accuracy

Word Parsing

Phrase parsing

Rule bases

Concept extraction

Are we AI?

Access Innovations, Inc.

We do not have

Neural Net
Bayesian statistics
Vector analysis
Co-Occurrence*

Yes, we fully qualify as an AI System

* Co-occurrence in our system is based on counts of occurrences

Summary

- Automated Indexing uses Artificial Intelligence techniques like ...
  - Rules
  - NLP
  - Machine Learning
  - Neural Nets
  - Latent Semantics
  - Co - Occurrence
  - Other Statistical Means
- So ... AI is AI
- Or there is not a clear enough definition so call it whatever you like
  - ...... Its all "marketing spin"

Thank you for your time

Marjorie Hlava

Access Innovations, Inc.

mhlava@accessinn.com

Taxonomy Thesaurus, Ontology

- **Taxonomy** – A collection of controlled vocabulary terms organized into a hierarchical structure. Each term in a taxonomy is in one or more parent/child (broader/narrower) relationship to other terms in the taxonomy.
  https://www.niso.org/publications/ansiniso-z3919-2005-r2010

- **Thesaurus** – A controlled vocabulary arranged in a known order and structured so that the various relationships among terms are displayed clearly and identified by standardized relationship indicators. Relationship indicators should be employed reciprocally. https://www.niso.org/publications/ansiniso-z3919-2005-r2010

- An **ontology** encompasses a representation, formal naming, and definition of the categories, properties, and relations between the concepts, data, and entities that substantiate one, many, or all [domains of discourse](#). More simply, an ontology is a way of showing the properties of a subject area and how they are related, by defining a set of concepts and categories that represent the subject.
  https://en.wikipedia.org/wiki/Ontology_(information_science)

Some key terms

- Rules engine
  - A collection of if/then statements that can execute a business function.
- Classification
  - An entity that can assigned a value. Can represent literally anything, so long as it can be named.
- Neurons
  - A mathematical representation of a human neuron in a neural network. Can have different "activation" parameters based on what type of neuron it is.
- Hyperparameters
  - The various tweakable configurations of a neural network, such as neuron count and activation types for the neurons.