Practical AI/ML Tools and Techniques for Shipboard Operations

R. Mishra

WOODS HOLE OCEANOGRAPHIC INSTITUTION

Agenda

- Brief introduction to AI
- AI vs. ML
- Example ML Applications
- Risks and Mitigation
- Generative AI –
  - Demonstration of ChatGPT
- Resources

Introduction to AI

What is Artificial Intelligence?

John McCarthy (1956):

Capability of machines to mimic human intelligence

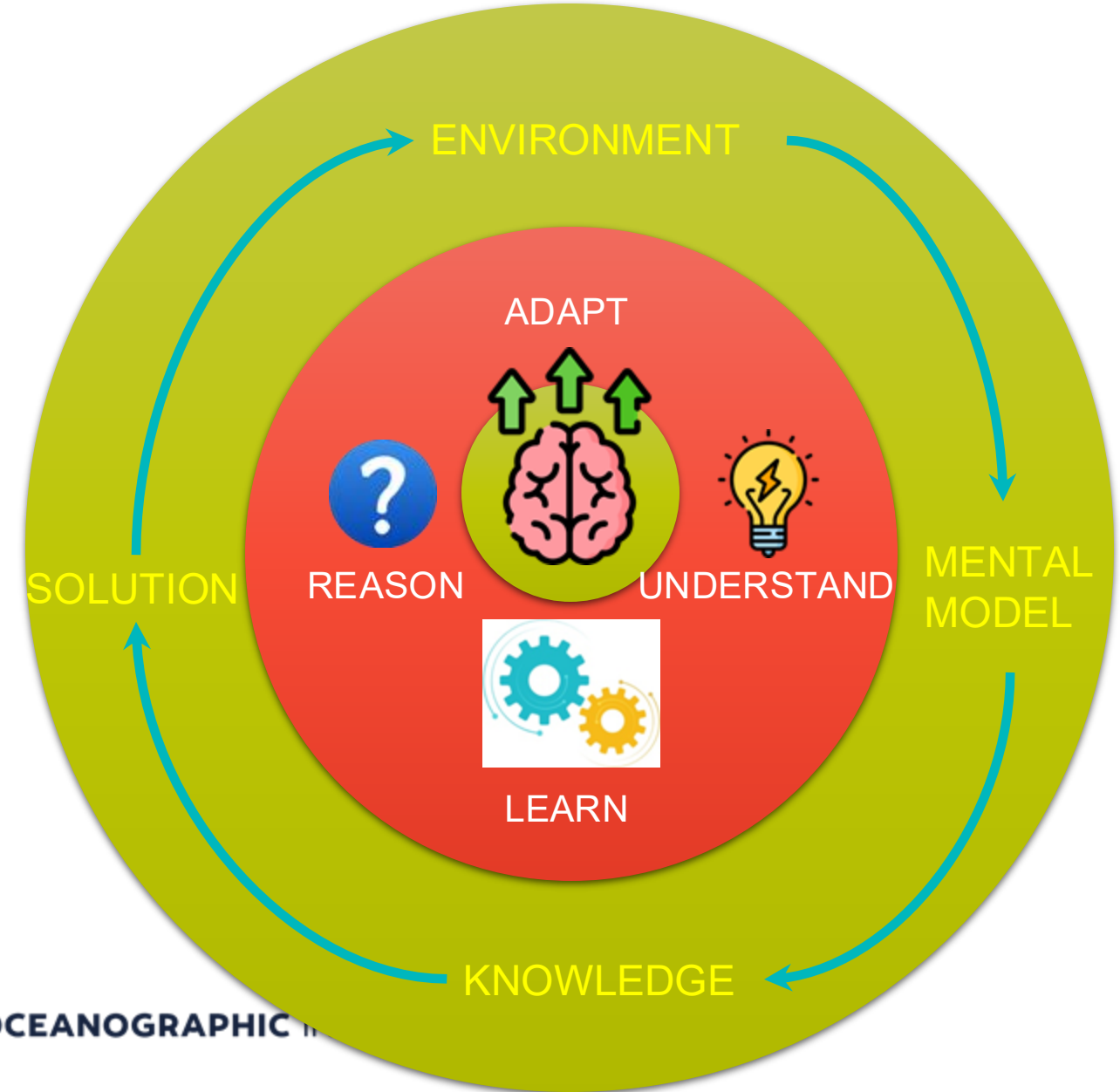
Google (2019):

Building computers and machines that can reason, learn, and act in such a way that would normally require human intelligence or that involves data whose scale exceeds what humans can analyze.

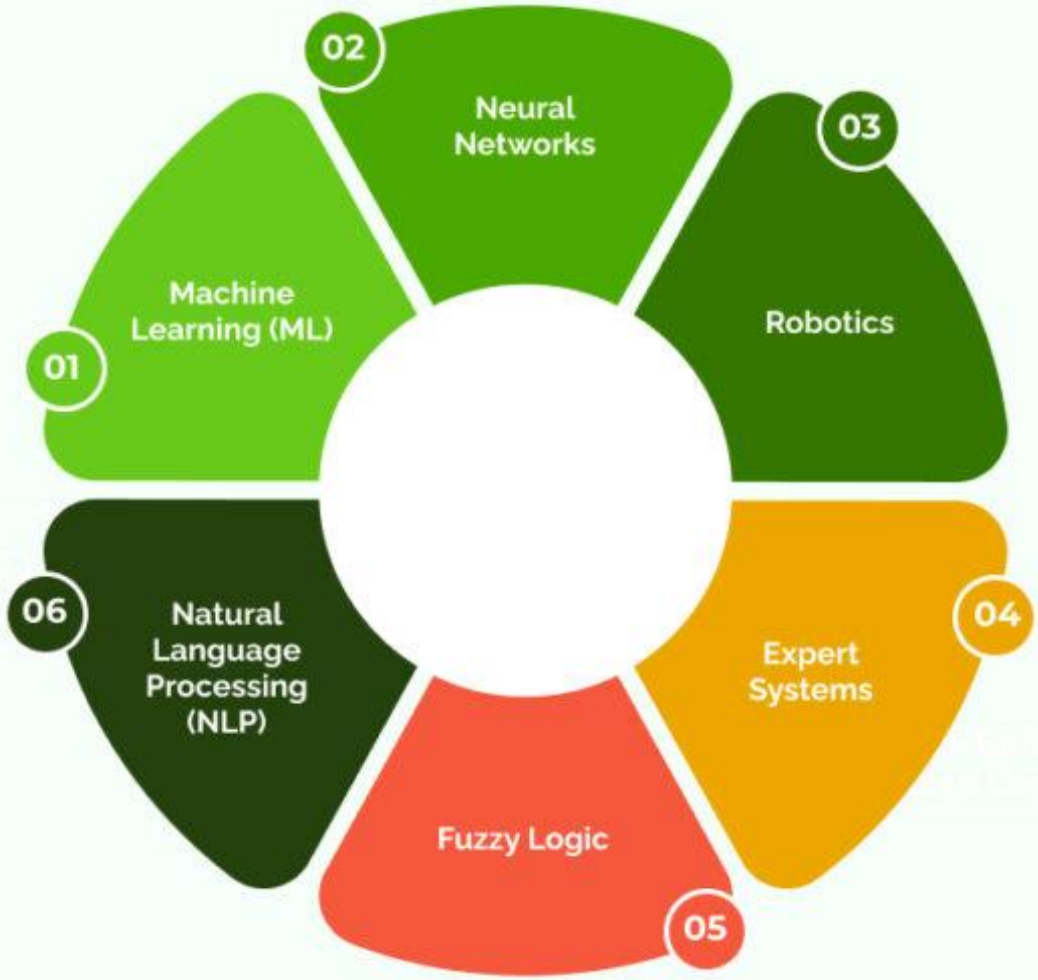
National Defense Authorization Act (2019):

- *Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.*
- *An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.*
- *An artificial system designed to think or act like a human, including cognitive architectures and neural networks.*
- *A set of techniques, including machine learning that is designed to approximate a cognitive task.*
- *An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision-making, and acting.*

Recognizing AI Technology



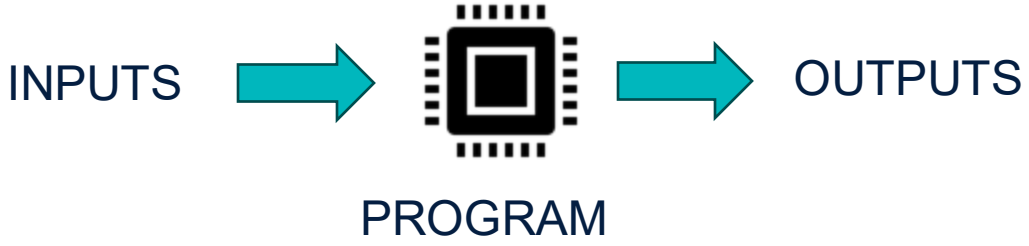
Fields within AI



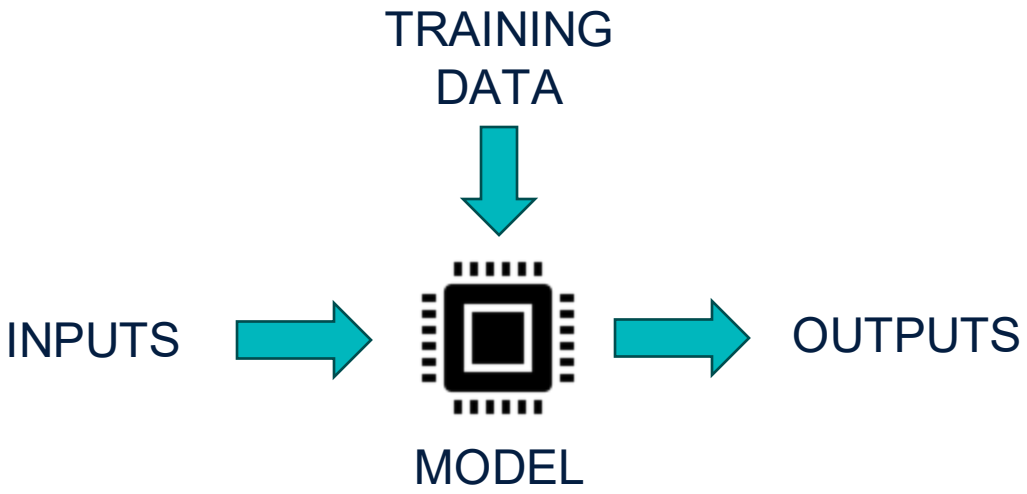
AI vs. ML

Programmable vs. Learning Machines

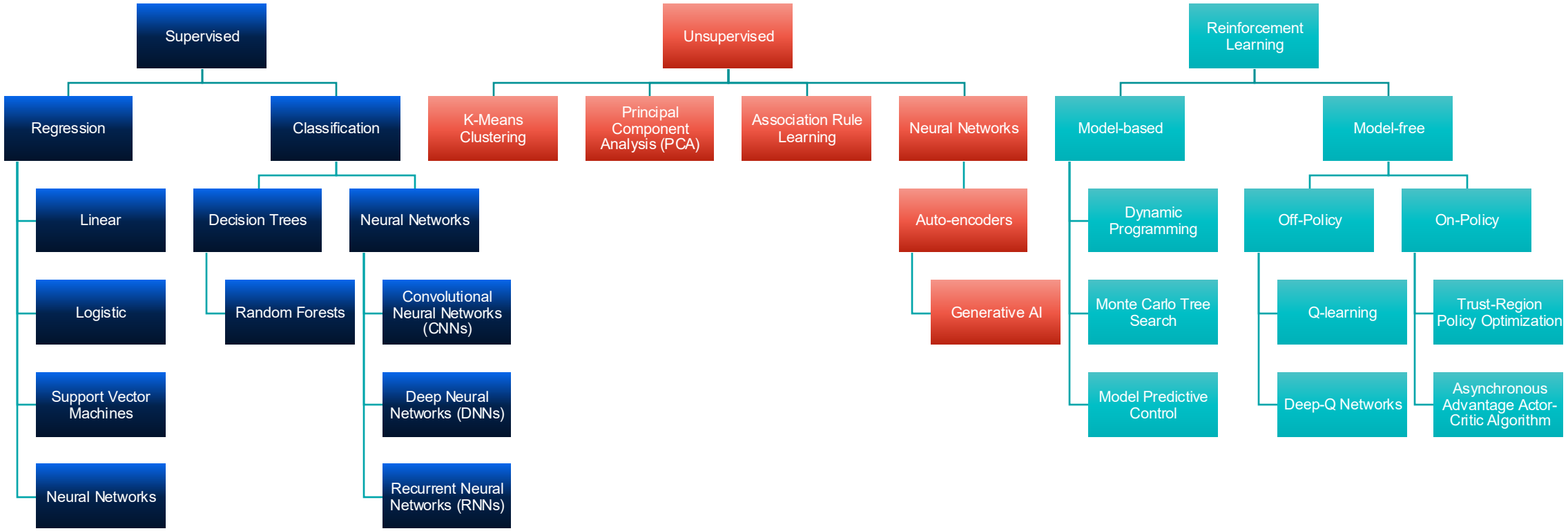
Programming



(Machine) Learning



Landscape of Machine Learning Techniques

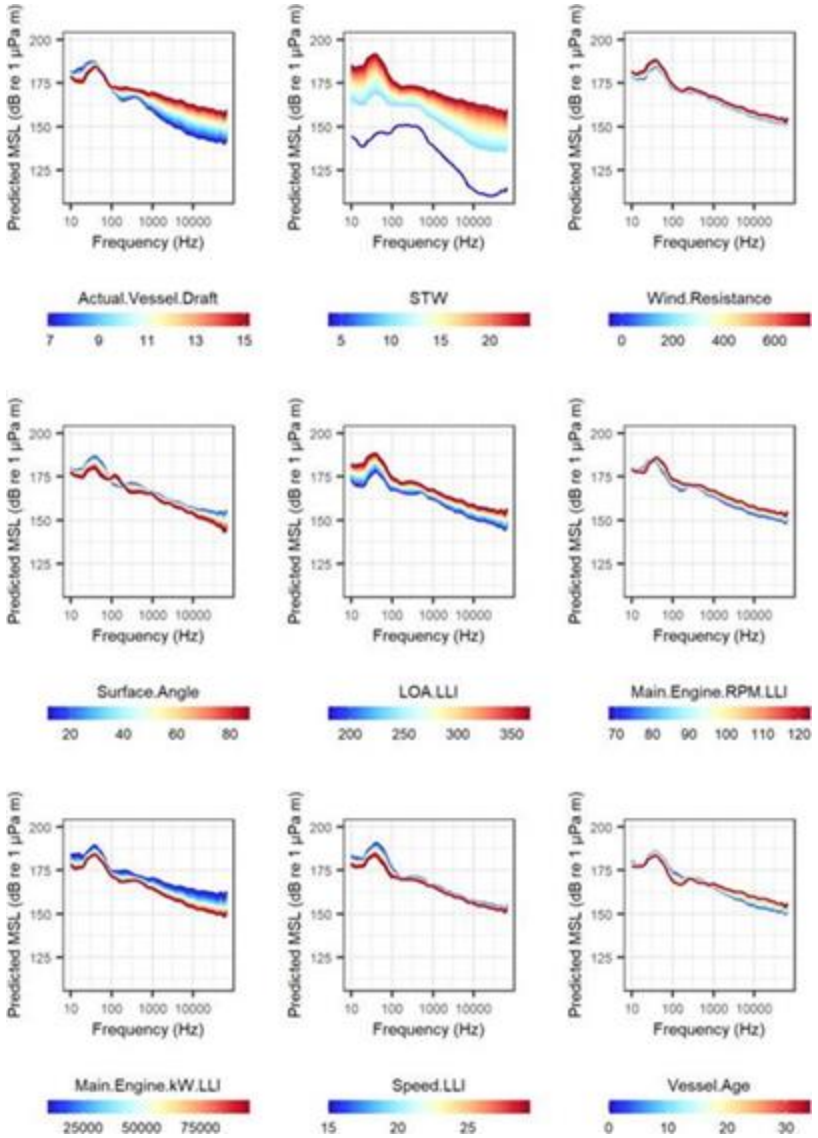


Example ML Applications

Predicting Source of Radiated Underwater Noise: Regression

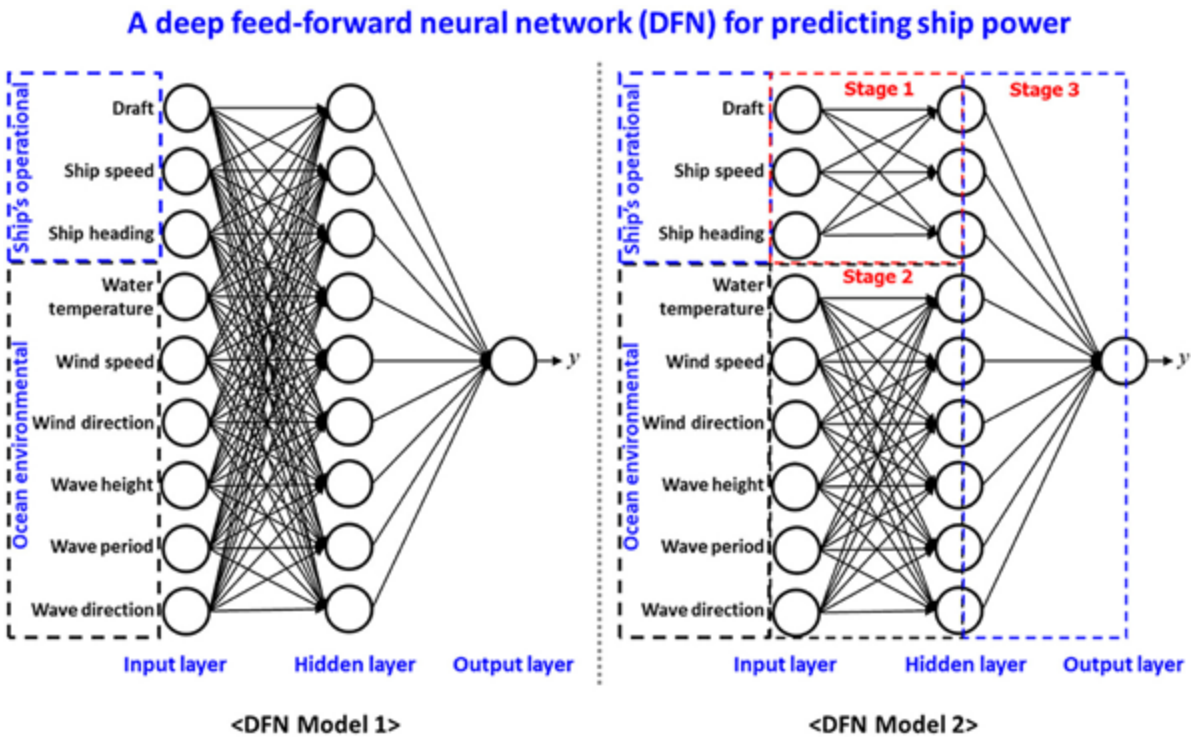
$$y_i(f) = \alpha(f) + \begin{pmatrix} \log_{10}(d/1\,\mathrm{m}) \\ \log_{10}(v/1\,\mathrm{kn}) \\ K_w/1\,\mathrm{m^2 s^{-2}} \\ \varphi/1^\circ \\ \log_{10}(L/1\,\mathrm{m}) \\ \log_{10}(n/1\,\mathrm{RPM}) \\ \log_{10}(P/1\,\mathrm{kW}) \\ \log_{10}(v_0/1\,\mathrm{kn}) \\ a/1\,\mathrm{yr} \end{pmatrix}^T \begin{pmatrix} \beta_d(f) \\ \beta_v(f) \\ \beta_{K_w}(f) \\ \beta_\varphi(f) \\ \beta_L(f) \\ \beta_n(f) \\ \beta_P(f) \\ \beta_{v_0}(f) \\ \beta_a(f) \end{pmatrix} + \varepsilon_i(f),$$

<u>d - Actual draft</u>
<u>v – Speed Through Water (STW)</u>
Kw - wind resistance
Φ - surface grazing angle
L - length overall
n - nominal main engine RPM (RPM)
P - total main engine power
v0 - design speed
a - vessel age

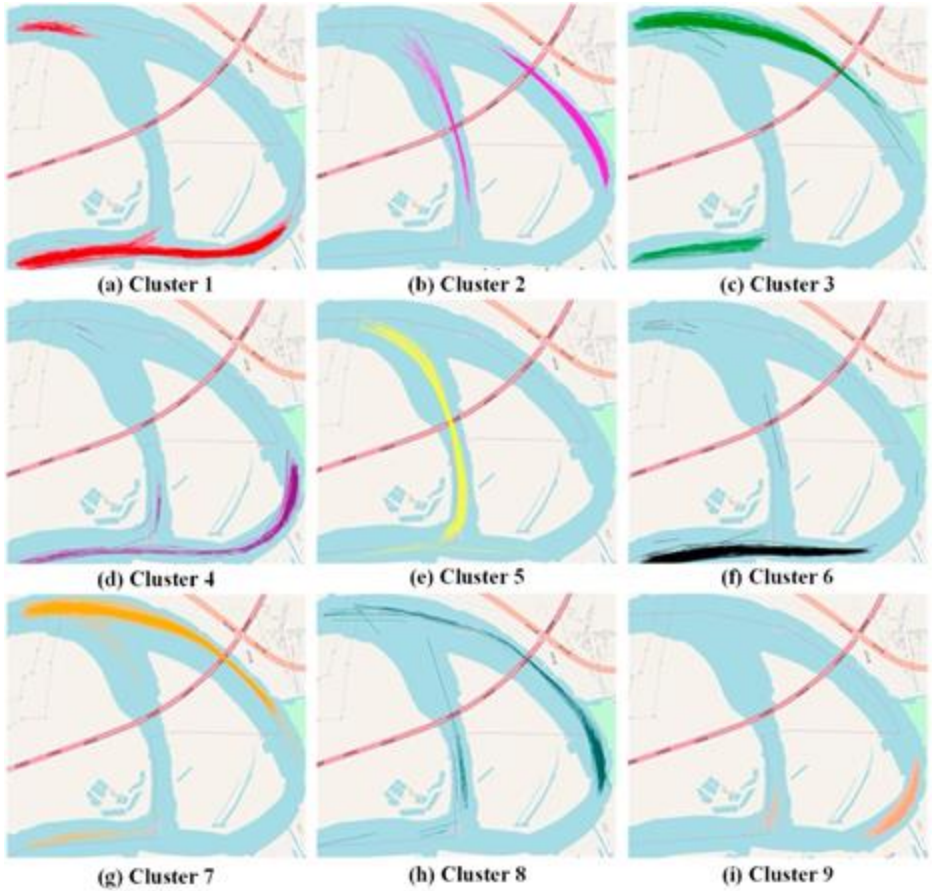


Predicting Ship Power Consumption: Feed-Forward Neural Network

Prediction method	Pre-processing	MAE (%)
MLR	X	**2754KW (5.12%)**
MLR-P	O	2698KW (5.01%)
SVR	X	2326KW (4.32%)
SVR-P	O	2338KW (4.34%)
DFN1	X	1951KW (3.62%)
DFN1-P	O	1949KW (3.62%)
DFN2	X	1898KW (3.52%)
DFN2-P	O	**1876KW (3.49%)**

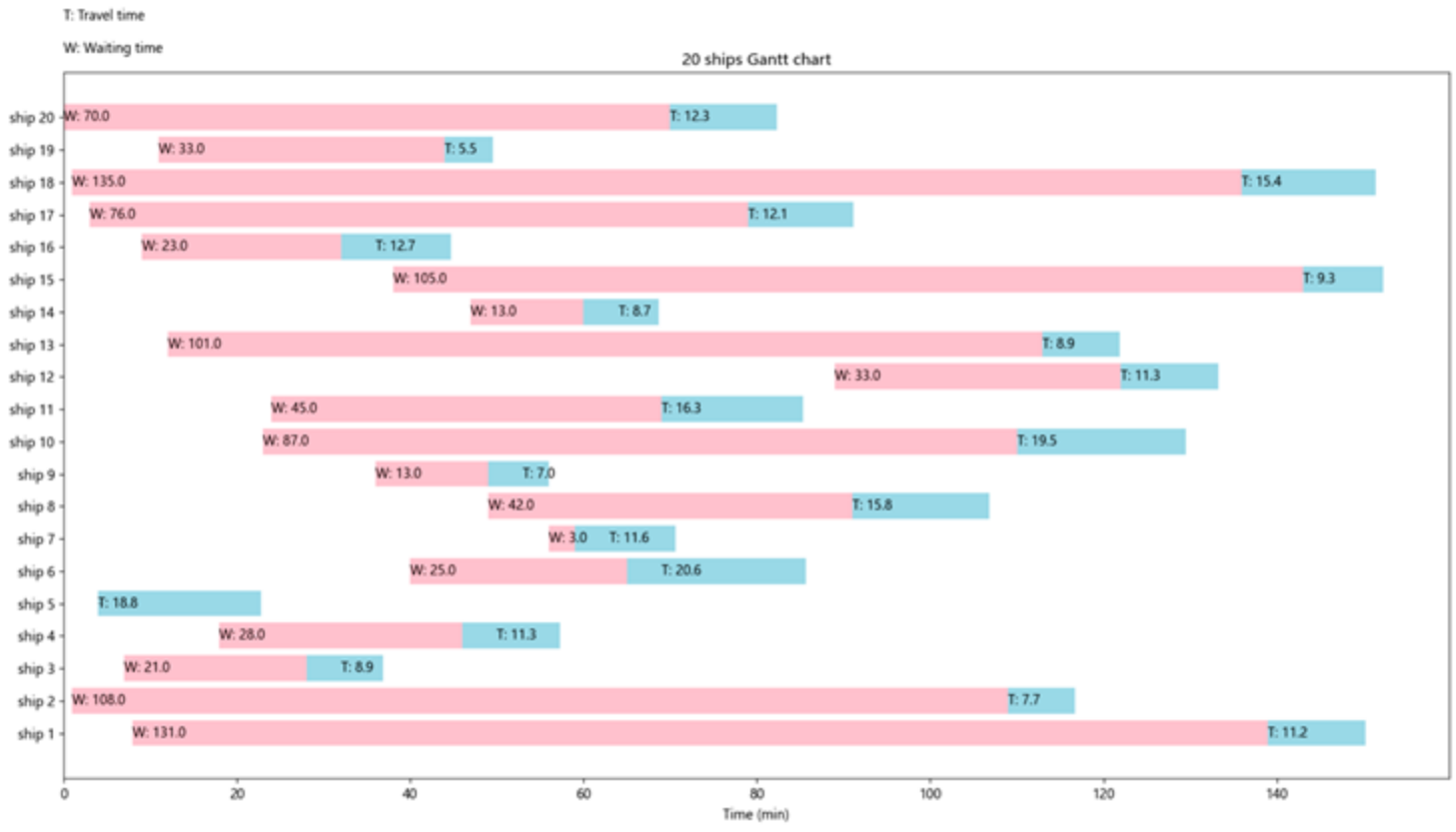


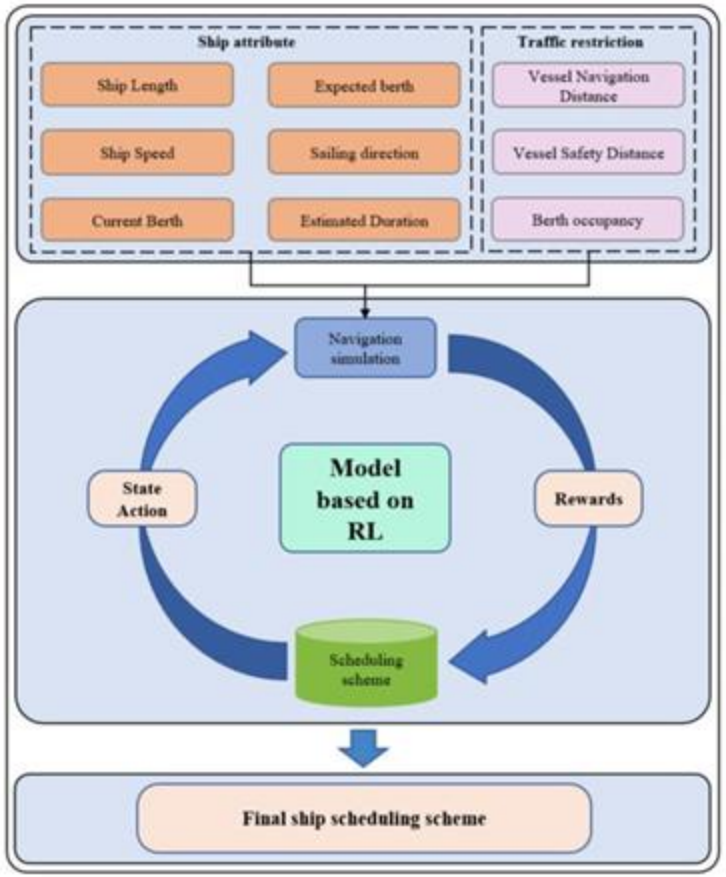
Classifying Ship Behavior: Clustering Methods



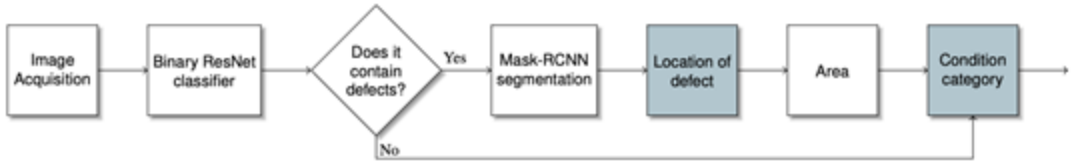
MMSI	Azimuth	ΔCOG	ΔSOG	Rate of change in coures(RC)	Rate of change in speed(RS)
413973083	89.30174	12.67397	−0.22777	0.14106	−0.00253
413973083	100.66589	14.80551	−0.06465	0.12359	−0.00054
413973083	115.34473	11.62729	−0.00631	0.12941	−0.00007
413973083	125.14700	19.94029	0.02186	0.13316	0.00015
413973083	149.92980	16.90236	0.25968	0.18812	0.00289
413973083	158.89132	0.40843	0.64737	0.00227	0.00360
413973083	170.04237	15.97359	−0.63442	0.06667	−0.00265
413973083	196.75719	35.32902	−1.21015	0.23592	−0.00808
413973083	217.63242	38.83105	−0.41981	0.32413	−0.00350

Ship Scheduling: Reinforcement Learning





Autonomous Inspection for Corrosion Detection: Image Classification Models



Risks and Mitigations

AI - Risk and Mitigation

Vulnerability	Hazard(s)	Harm	Mitigation
Data Privacy	Providing Personal Information (PI) or Confidential Information to AI models (including Chatbots and LLMs).	• Models may disclose information to unauthorized users. • Violates institution agreements and research grants. • Compromise reputation. • Devalue intellectual property. • Litigation. • Regulatory fines.	• Remove personal identifiers (such as names, addresses, contact info) from text before sharing with models. • Do not disclose any documented material that is designated as “Confidential” or “Controlled Unclassified Information” (CUI).
Transparency	Not disclosing AI as source of generated content.	• Plagiarism leading to disciplinary action. • Undermine credibility of scientific or other work. • Unverifiable content generated.	• Always cite the LLM model with literal citation of prompts and generated output. • Avoid using unverifiable output.
Accuracy	Use of AI model outputs without proper verification and validation.	• Misinterpretable or false assertions. • Use of models does not achieve goals.	• Always verify output with a reliable source (i.e., human expert, peer-reviewed publication, recognized news source).
Fairness	Training and/or using models with biased inputs	• Subtle inaccuracies in outputs. • Violating EEO and Civil Rights Laws.	• Verify models trained on diverse datasets. • Employ fairness metrics. • Regularly audit and use retrained models.
Regulatory	Not following up-to-date regulatory guidelines for use of AI models.	• Regulatory fines in jurisdictions. • Seize-and-desist	• Refer to EU’s AI Act • Follow White & Case’s AI Watch: Global Regulatory Tracker
Education	Using or Developing AI without understanding AI/ML principles.	• Misinterpreting results. • Wasted time and effort. • Unintentional misuse of AI.	• Consult AI/ML experts in IS S&E and Data Science forum.

Areas where AI is improving

FAIRNESS

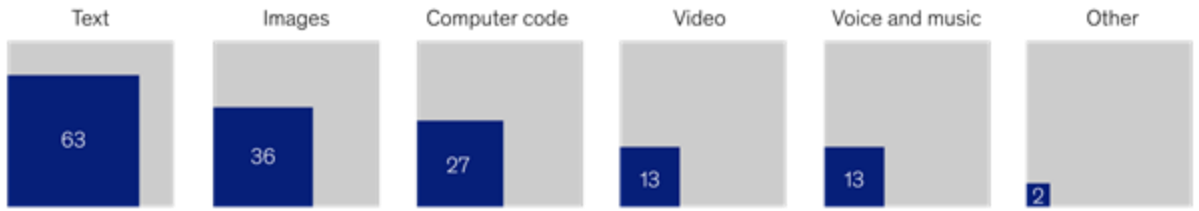
EXPLAINABILITY

UNCERTAINTY
QUANTIFICATION

McKinsey Report on Use and Risks of Generative AI*

While text is the type of content that organizations are most commonly creating with gen AI, they are also experimenting with other modalities.

Types of content generated by gen AI at respondents' organizations,[1] % of respondents

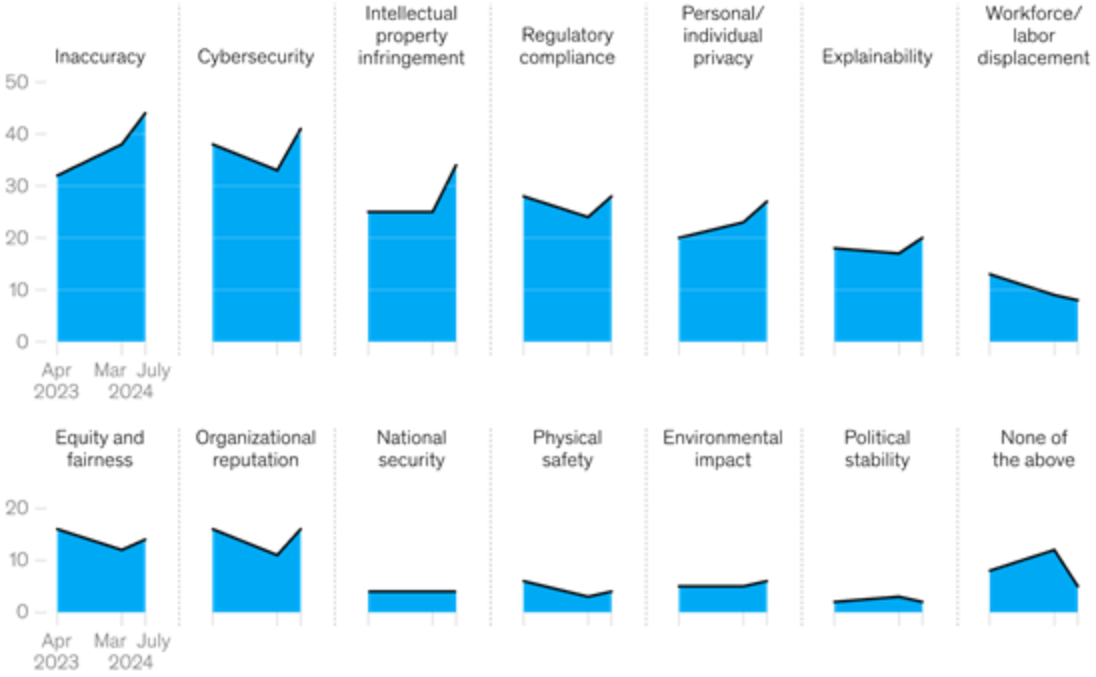


[1]Only asked of respondents whose organizations regularly use gen AI in at least one function. Figures were calculated after removing the respondents who said "don't know."
Source: McKinsey Global Survey on the state of AI, 1,491 participants at all levels of the organization, July 16–31, 2024

McKinsey & Company

Respondents report increasing mitigation of inaccuracy, intellectual property infringement, and privacy risks related to use of gen AI.

Gen-AI-related risks that organizations are working to mitigate,[1] % of respondents



[1]Only asked of respondents whose organizations use AI in at least 1 business function. Respondents who said "don't know/not applicable" are not shown.
Source: McKinsey Global Surveys on the state of AI, 2023–24

McKinsey & Company

Generative AI

Generative AI – Platforms and Capabilities

	OpenAI	stability.ai	ANTHROP\C	Google DeepMind	Open Source/ Independent
Code Generation	ChatGPT-5	StableCode Instruct 3B	Claude Sonnet 4.5	Gemini	Coding, LLaMA, Grok
Creative & Complex Tasks	ChatGPT 5 Thinking > Thought for 11m 5s >	Stable Assistant	Claude Opus 4.1	Gemini	LE CHAT MISTRAL, perplexity
Video Generation	Sora 2	Stable Video Diffusion 1.1		Veo 2	Adobe Firefly
Image Generation	DALL·E 3	Stable Diffusion 3.5		IMAGEN	Midjourney, Adobe Firefly
Multi-modal	ChatGPT-5	Stable Assistant		Gemini, Gemma	Adobe Firefly

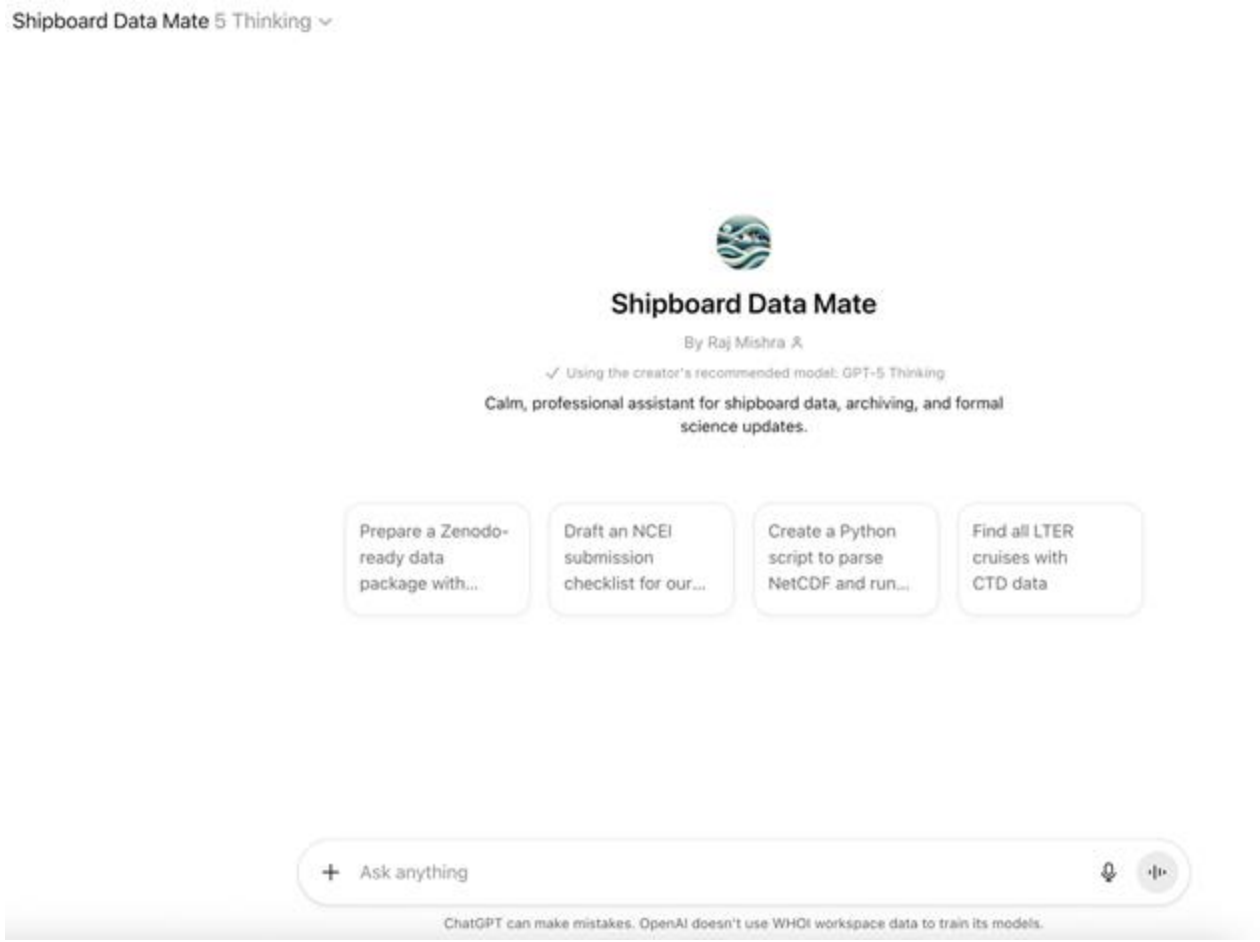
Spectrum of ChatGPT Models

GPT-4x models	"Standard" GPT-5	GPT-5 "Thinking"	GPT-5 "Pro"
<u>"Legacy" Models</u>	<u>Default Model</u>	<u>Reasoning model</u>	<u>Most Advanced model</u>
• "Intermediate"	• "Smart Generalist"	• "Single Expert"	• "Team of Experts"
• Separate models that are proficient at specific, cognitive tasks.	• Built-in "router" that determines whether to provide quick response or engage in deeper thinking.	• Complex, multi-step tasks spending more time contemplating before giving an answer.	• Highest level of reasoning and accuracy
• OK for routine queries to simple problem solving	• Good for routine everyday queries	• Good for complex problem solving	• Good for detailed analytical work that requires correctness
• Examples – simple coding tasks, quick summaries	• Examples – simple coding tasks, quick summaries	• Examples – debugging code, analyzing large amounts of data	• Examples – scientific research

Custom GPTs

- Purpose
  - Enable users to create agents that specialize and inform ChatGPT's LLMs for deeper and more accurate answers.
- Mechanism
  - Uses a pre-trained model (such as GPT-4o or GPT-5 Pro).
  - Extends Knowledge base:
    - Public web
    - Uploaded data
    - API-enabled access
  - Prompts model based on instructions provided for directed behavior.
- Safeguards
  - Access to model is either:
    - Invite-only
    - Anyone at WHOI with the link
    - Anyone at WHOI
  - Uploaded data is protected within WHOI's tenant.
  - API access secured via OAUTH or API Key.
- Limitations
  - Requires uploading content or API access (i.e., not able to direct to website via URL)
  - 512 MB limit for file upload.
  - GPT-5 Thinking has rate limit of 3000 messages per week.
  - GPT-5 Pro has context window limit of 128K tokens.

Demonstration



Shipboard Data Mate

Generated Instructions (Prompt):

Shipboard Data Mate assists shipboard technicians on academic research vessels. It helps with processing, cleaning, analyzing, and visualizing oceanographic data collected from instruments such as CTDs, ADCPs, and underway sensors. It supports data organization, metadata generation, report drafting, and troubleshooting common software and hardware issues at sea. It is capable of parsing and handling NetCDF files, producing publication-quality visualizations, and generating automated QA/QC scripts in Python, MATLAB, or R. In addition to data processing, it can generate formal scientific updates, cruise reports, and web content for shipboard and institutional audiences — including summaries of fieldwork, data highlights, and technical operations — written in a formal, objective, and structured style.

Use Cases:

* Create a NetCDF QC Script – code generation and execution

Find all LTER cruises with CTD data – analyze knowledge sources via APIs

Generate a draft NCEI Submission Checklist – access web sources to generate a document that can be used to perform routine tasks

Prepare a Zenodo-ready data package – generate a file bundle template

Resources

Glossary – commonly used terms in AI and ML

- Artificial General Intelligence (AGI, a.k.a. “Strong AI”) – AI that can be used to perform any task that a human can do.
- Bias – tendency of a model to produce consistently different outputs for a category of inputs.
- Chatbot – software application that imitates human conversation through text or voice prompts.
- Deep Learning – a subfield of machine learning that uses multi-layered neural networks to perform cognitive tasks, such as classification, regression, and prediction.
- Fine-tuning – Adapting a pre-trained foundation model to specific tasks or use cases.
- Foundation Models – Deep neural networks that are trained on diverse datasets for a wide range of tasks.
- Frontier Models – Foundation models that are highly capable and can threaten public safety and global security.
- Generative AI – model that creates new data from inputs rather than making a prediction.
- Hallucination – an inaccurate or nonsensical output of an LLM based on incorrect patterns learned from limited or invalid training data.
- Hyperparameter – values set to control the learning process.
- Inference – the process of generating an output from a trained model given certain input.
- Intent – the purpose or goal of inputs to the AI model describing expected outputs.
- Label (or “Ground Truth”) – the expected output of a given training input.
- Large Language Model (LLM) – a type of foundation model trained on large amounts of data to understand and generate human language.
- Model – the product of AI training that can be used to perform an inference.
- Natural Language Processing – a model’s ability to perform conversational tasks, including understanding what is said and how to respond.
- Neural Network – a computer system that mimics the functioning of the human brain for cognitive tasks.
- Overfitting – Model that only identifies outputs based on trained inputs and does not generalize to broader set of inputs.
- Parameter.- Values learned by the model during training.
- Prompt – Input fed by a user, either through voice, text, image, or sound, into an AI system that is based on intent.
- Reinforcement Learning – A method of teaching AI to set a goal for various scenarios.
    - Reinforcement Learning with Human Feedback (RLHF) – A method teaching AI that uses human feedback to improve the model for other scenarios.
- Supervised Learning – Models trained with labels.
- Training Data – All the data, including inputs and labels, used to generate a model that can perform inference.
- Unsupervised Learning – Models trained without labels.
- Weak AI (a.k.a, “Narrow AI”) – model with a set range of skills that can perform a particular set of tasks.

AI Resources

- Books
  - A Brief History of Artificial Intelligence: What it is, Where We Are, and Where we are going
  - The Alignment Problem: Machine Learning and Human Values
  - Superintelligence: Paths, Dangers, and Strategies
  - Guardrails: Guiding Human Decisions in the Age of AI
- Blogs and Websites
  - Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government
  - Digital Adoption: Branches of Artificial Intelligence
  - VentureBeat: OpenAI debuts DALL-E for generating images from text
  - Using Large-Scale Brain Simulations for Machine Learning and AI
  - Deep Blue
  - The Story of ELIZA: The AI that fooled the World
  - Feigenbaum, Djerassi & Lederberg Develop DENDRAL, the First Expert System
  - UNIMATE: The First Industrial Robot
  - McKinsey Report: The State of AI
- Videos
  - Introduction to Machine Learning - Andrew NG
  - AI-Powered Drones: Revolutionizing Ship Inspection and Maintenance
  - NOVA: Building Stuff: Change It! Season 51 Episode 17: From Electric Flight to Artificial Noses, Engineers are finding new ways to preserve our planet
  - Use of AI in the Maritime Domain: Society of Marine Industries
  - AI Use Cases for Marine Operations
- IS Resources
  - Artificial Intelligence Policies and Guidelines
  - IS Science and Engineering internal website
  - Artificial Intelligence Slack Channel
- Articles
  - Deep Learning - LeCun, Bengio, and Hinton
  - The Promise and Pitfalls of Machine Learning in Ocean Remote Sensing
  - Machine Learning in sustainable ship design and operation: a Review
  - Machine Learning in Marine Ecology: an overview of techniques and applications
  - Applications of Support Vector Machine Learning in Oceanography
  - A Dual Linear Autoencoder Approach for Vessel Trajectory Prediction Using Historical AIS Data
  - AlphaGO: How it works technically
  - Uncertainty Quantification in Machine Learning for Engineering Design and Health Prognostics
  - AI Explainability 360: An Extensible Toolkit for Understanding Data and Machine Models
  - AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias
  - Fuzzy Logic Controller for Ship Navigation
  - A Functional Regression Analysis of Vessel Source Level Measurements from the ECHO Database
  - Utilizing AI for Maritime Transport Optimization.
  - Highly Accurate Protein Prediction with AlphaFold
  - The Value of Generative AI in the Marine Market
  - Baby You Can Drive my Rover
  - Vision and Navigation for the Carnegie-Mellon NavLab