AMERICAN SOCIETY OF SAFETY PROFESSIONALS
New Mexico Chapter

NMSHMM

Artificial Intelligence and System Safety:

A general discussion on the incorporation of system safety into the development of AI systems as well as the use of AI to perform system safety

Timothy Stirrup CSP CHMM REM & Kelsey Forde CIH CSP CHMM
Parvati Consulting & Government Services
Joint Technical Meeting - Thursday, October 10th 2024

1

Timothy Stirrup, CSP CHMM REM

- Partner at Parvati Consulting, LLC & Parvati Government Services Inc
- Principal EHS Professional & System Safety Engineer
- B.S., Biology and B.S. Chemistry from the New Mexico Institute of Mining and Technology (NM Tech)
- Over 35-years of experience as an EHS professional including performing and guiding hazards analyses that adhere to the principles in the "Redbook."
- Primary responsibilities and areas of expertise are centered around establishing the framework for clients in Hazard Analysis within a diverse set of industrial facilities including accelerators, semiconductor facilities, R&D laboratories, and dynamic energetic materials testing facilities.
- Respected for his ability to create highly functional teams and solve diverse, complex problems.
- Provides direct input to Line/Executive Management and Regulators to ensure solutions for continued organizational success.
- President of the New Mexico Society for Hazardous Materials Managers Board of Directors, Past-President for the New Mexico Chapter of the American Society of Safety Professionals, Vice President of System Safety Society New Mexico Chapter, and Operational Vice President of Marketing for International Society of System Safety
- Awards/Recognition: 2020 International Society of System Safety Educator of the Year Nominee; 2022 Alliance of Hazardous Materials Professionals Distinguished Lecturer (Inaugural Class)

2

Kelsey Forde, CIH CSP CHMM

- Owner of Parvati Consulting LLC & Parvati Government Services Inc
- Principal EHS Professional & System Safety Engineer
- M.S. Environment, Health, and Safety from University of Minnesota, Duluth
- B.S. Cellular Biology (Chemistry & Pre-Pharmacy Minor) from University of Minnesota, Duluth
- Nearly 20-yrs experience as an EHS professional including performing and guiding hazards analyses that adhere to the principles in the Redbook
- Primary responsibilities and areas of expertise are centered around the identification of workplace hazards and development of consequence analysis associated with hazard analysis, safety assessments, primary hazard screens, readiness reviews, and compliance auditing techniques for a variety of clients including the DOE, commercial, industrial, and private industry clients.
- National Director & Chair for the Alliance of Hazardous Materials Professionals, Past-President and Current President for the New Mexico Chapter of the American Society of Safety Professionals, Past-President and Director at Large for the New Mexico Society of Hazardous Materials Managers, President of System Safety Society New Mexico Chapter, and historically served two consecutive mayoral appointed terms on the Albuquerque-Bernalillo County Joint Air Quality Control Board.
- Awards/Recognition: 2022 Energy Facilities Contractor Group (EFCOG) Teamwork Award recipient for contributions to the update of DOE-HDBK-1163-2020, Integration of Hazard Analyses; 2022 ASSP Safety Professional of the Year (SPY) Award by the New Mexico Chapter; 2022 AHMP Distinguished Lecturer (Inaugural Class)

3

4

Video Clip – Terminator 3: Rise of the Machines

- 2003 science fiction action film, the third installment in the Terminator franchise and a sequel to Terminator 2: Judgment Day (1991)
- Directed by Jonathan Mostow and stars Arnold Schwarzenegger, Nick Stahl, Claire Danes, and Kristanna Loken.
- Plot: The malevolent artificial intelligence Skynet sends a T-X (Loken)—a highly advanced Terminator—back in time to ensure the rise of machines by killing top members of the future human resistance as John Connor's (Stahl) location is unknown. The resistance sends back a reprogrammed T-850 (Schwarzenegger) to protect John and Kate (Danes).
- You Tube Video Clip "Skynet Takes Over" https://youtu.be/_Wlsd9mljiU?si=qX6CGxrsE5KXznsJ (4:21)

Ref. – Wikipedia; https://en.wikipedia.org/wiki/Terminator_3:_Rise_of_the_Machines

5

Artificial Intelligence - General Discussion

- Artificial intelligence (AI) is in our every day lives.
- AI is a general-purpose technology across industries including transportation, healthcare, financial services, and industrial engineering.
- An understanding of AI is imperative for the system safety practitioner for understanding of systems as well as a supporting tool.

6

Artificial Intelligence - General Discussion

- AI is perceived intelligence exhibited by machines, particularly computer systems, as opposed to the natural intelligence of living beings
- AI is a field of research in computer science that develops and studies methods and software which enable machines to perceive their environment and uses learning and intelligence to take actions that maximize their chances of achieving defined, peer reviewed goals
- Development of AI systems must consider system safety aspects reflecting standard practices with system safety engineering

7

Artificial Intelligence - Direction

- Narrow AI - Designed for a specific task (like language translation)
- General AI – Ability to perform any intellectual task that a human can do

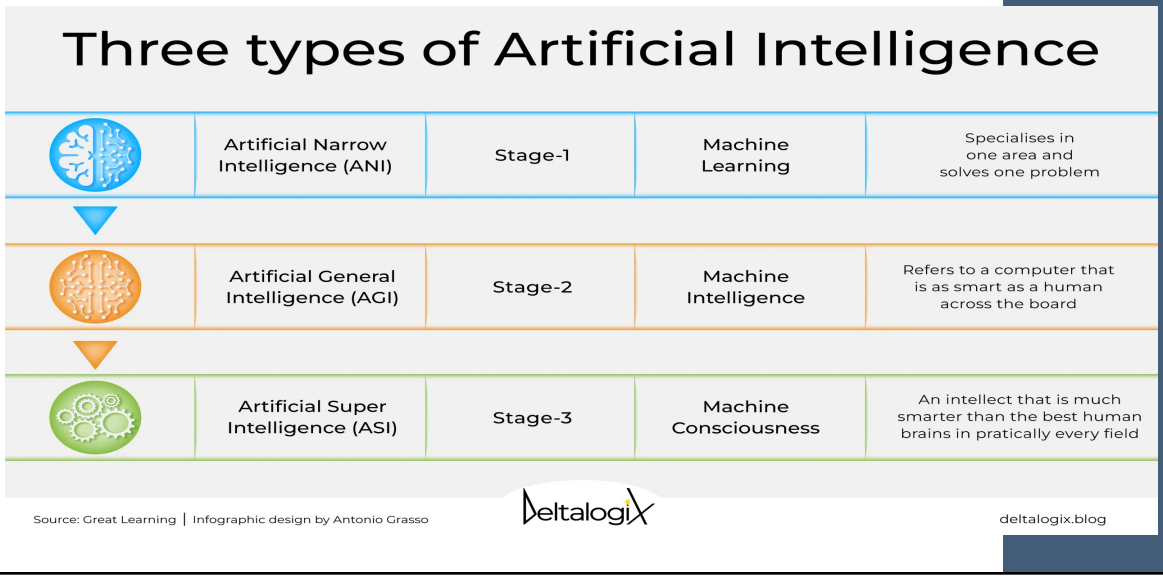
8

Artificial Intelligence – Defined by AI

- AI refers to the simulation of human intelligence processes by machines, particularly computer systems:
  - Learning (the acquisition of information and rules for using it)
  - Reasoning (using rules to reach approximate or definite conclusions)
  - Self-correction.
- Specific applications of AI include:
  - expert systems
  - speech recognition & natural language processing (NLP)
  - machine vision.

9

Artificial Intelligence – Types



10

4 Types of AI Based on Functionality

- **Reactive Machine AI** - Reactive machines are AI systems with no memory and are designed to perform a very specific task. Since they can't recollect previous outcomes or decisions, they only work with presently available data.
- **Limited Memory AI** - AI that can recall past events and outcomes and monitor specific objects or situations over time. Limited Memory AI can use past- and present
- **Theory of Mind (Emotion) AI** - AI with Theory of Mind functionality would understand the thoughts and emotions of other entities. This understanding can affect how the AI interacts with those around them
- **Self Aware AI** - Self-Aware AI is a kind of functional AI class for applications that would possess super AI capabilities

11

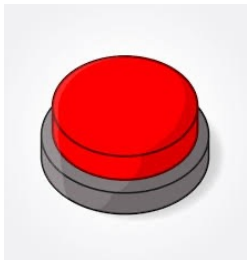
Development of AI in Systems

- System Safety Practitioners should be involved in development of systems with AI
- Application of standard failure analysis and identification of controls

12

Negative Aspects of AI

- Assume AI is Perfect
- Risk of over-reliance of AI in systems
- System safety professionals become more obsolete
- Over selling AI solutions in absence of human reviews
- AI systems should always have a "**Big Red Button**"



13

AI Is Not...

- **Human**: AI lacks emotions, consciousness, and personal experiences
- **Perfect**: AI can make mistakes and may provide incorrect or biased outputs
- **Self-aware**: AI does not have self-awareness or understanding of its existence

14

AI Is Not...

- **Independent**: AI requires human input and guidance for development and operation
- **Magic**: AI is based on algorithms and data, not on any mystical or unexplained forces
- **Omniscient:** AI's knowledge is limited to the data it has been trained on and cannot understand or infer beyond that
- **Creative**: While AI can generate creative outputs, it does not possess true creativity or original thought in the human sense

15

Notable AI Applications – Examples

- AI has several notable applications in system safety engineering, enhancing the safety, reliability, and efficiency of various systems.
- These applications demonstrate how AI can significantly contribute to improving the safety and reliability of engineered systems across various industries, including aerospace, automotive, manufacturing, and healthcare

16

Notable AI Applications – Examples

- Predictive Maintenance: AI algorithms analyze data to predict equipment failures enabling preemptive maintenance and reducing downtime and safety risks.
- Fault Detection and Diagnosis: Machine learning models can detect anomalies in system behavior that may indicate faults or potential safety issues.
- Risk Assessment: AI tools can assess and quantify risks in complex systems by analyzing historical data, environmental conditions, and operational parameters.
- Safety Monitoring: AI systems continuously monitor operations, identifying unsafe conditions or behaviors and triggering alerts or automated responses.

17

Notable AI Applications – Examples

- Human Factors and Ergonomics: AI analyzes human interaction with systems to identify potential safety risks related to human error and suggest improvements.
- Simulation and Modeling: AI enhances simulation tools used in system safety engineering by providing more accurate models of complex systems,
- Decision Support Systems: AI-driven decision support systems assist engineers in making safety-critical decisions by providing complex analyses.
- Automated Safety Analysis: AI can automate parts of the safety analysis process, such as hazard analysis and FMEA, improving efficiency and consistency.

18

Advice to the System Safety Professional Addressing AI

The System Safety Profession should address AI by developing comprehensive strategies that integrate AI technologies while ensuring safety, reliability, and ethical standards. Consider:

- Education and Training: Skill Development; Continuous Learning
- Integration with Existing Processes: Incorporate AI Tools; Augment Human Expertise
- Regulatory and Ethical Standards: Develop Guidelines; Ensure Compliance
- Risk Management: Identify AI-Specific Risks; Develop mitigation Strategies
- The Transparency & Explainability: Understandable Models; Explainability Tools

19

Advice to the System Safety Professional Addressing AI

By adopting these strategies, the System Safety Profession can effectively leverage AI to enhance safety while managing its associated risks and ethical implications.

- Collaboration and Sharing Best Practices: Foster Industry Collaboration; Form Interdisciplinary Teams
- Robust Testing and Validation: Simulation and Testing; Implement Continuous Monitoring
- Ethical Considerations: Bias and Fairness; Establish Accountability:

20

Potential Benefits of AI for System Safety

- AI can enhance system safety by improving predictive maintenance, risk assessment, and decision-making processes
- AI can analyze vast amounts of data quickly to identify potential safety hazards or risks in a system.
- AI-powered automation can help in reducing human errors and improving the overall safety of operations
- AI can analyze video real time to identify potential human and equipment interactions

21

Challenges of AI for System Safety

- AI systems can be susceptible to biases in data, which could lead to incorrect decisions affecting safety
- The complexity of AI algorithms can make it challenging to understand and validate their safety performance in all scenarios
- AI systems may exhibit unpredictable behaviors or make decisions that are difficult to interpret, leading to safety concerns
- AI development without incorporation of system safety

22

Safety Assurance in AI Systems

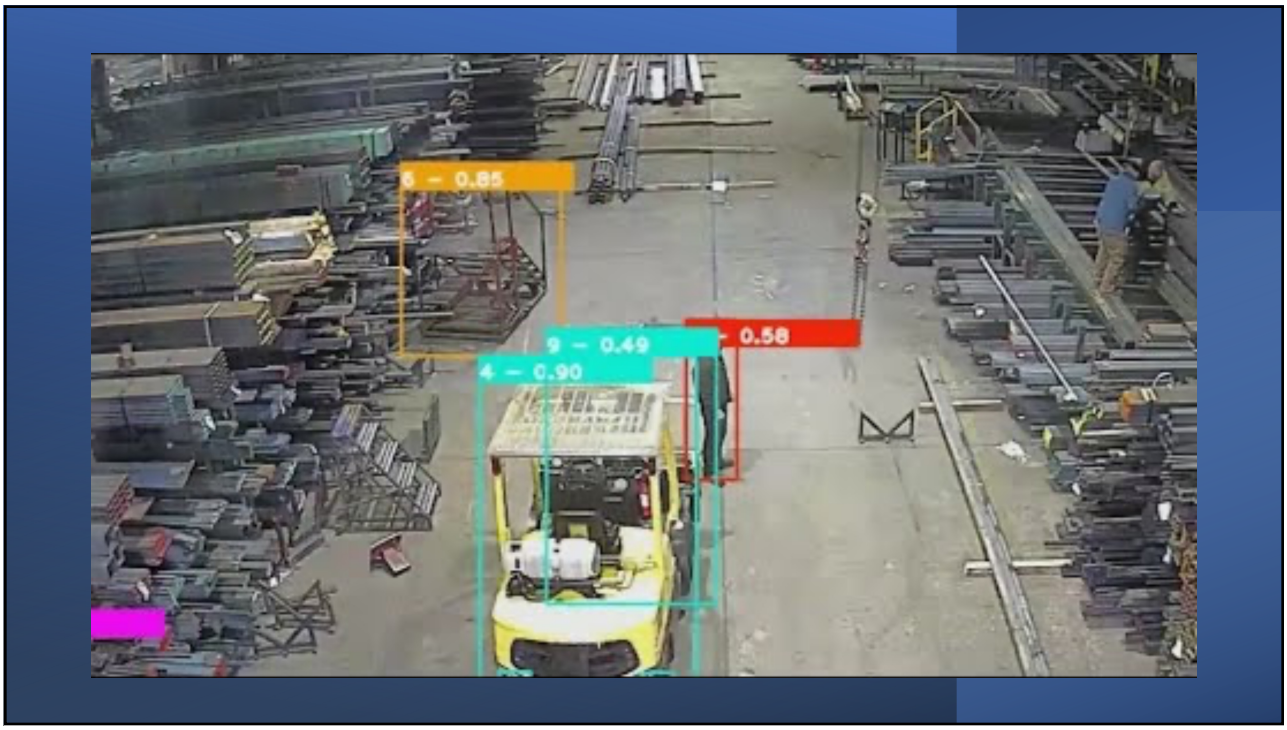
- Ensuring the safety of AI systems requires a combination of technical validation, regulatory compliance, and ethical considerations
- Safety assurance processes for AI systems should include rigorous testing, validation, and monitoring to identify and mitigate risks
- The development of AI safety standards and regulations is essential to ensure that AI systems meet acceptable safety levels

23

Human-AI Collaboration for System Safety

- Human oversight and collaboration are crucial for ensuring the safe integration of AI into systems
- Humans play a key role in interpreting AI outputs, challenging AI decisions, and taking control in case of unexpected situations
- Training programs for users and operators can help in developing the necessary skills to work effectively with AI systems for safety-critical applications

24



25

Video Clip: Real World Example – AI for Workplace Safety

Broadcast, Fox News 11, Good Day L.A.

- May 13, 2024
- "More companies are turning to AI to improve workplace safety"
- https://youtu.be/y_ARNq6MaT0?si=kTXQ7-p92nT2iBIX (1:53)

26

Summary Idea for AI & System Safety

AI offers significant opportunities to improve system safety but also presents unique challenges that need to be addressed through a combination of technical, regulatory, and human factors considerations to ensure the safe and responsible deployment of AI technologies in various systems. AI will not replace humans but will make our jobs more efficient.

27

Follow Up with Parvati…

Kelsey Forde CIH CSP CHMM
- KLForde@ParvatiCorp.com
- (505) 967-8917

Timothy Stirrup CSP CHMM REM
- TSSTirrup@ParvatiCorp.com
- (505) 980-3743

www.ParvatiCorp.com

- **Traditional ES&H/IH Services**
- **Facility/Worker Safety**
- **Hazard Evaluation Training**
  - **"Redbook" Overview**
  - **What-If/Checklist**
  - **Failure Modes & Effects Analysis**
  - **Hazard & Operability Analysis**
  - **Layer of Protection Analysis (LOPA)**
  - **Risk Analysis**
  - **Inherent Safety Reviews**
  - **STAMP/STPA**
- **Process Hazards Analysis**
- **Compliance Auditing**
- **Facilitate Hazard Evaluations**
- **Peer Review PHA (HI + HE)**

Parvati Consulting & Government Services
…woman-owned small businesses…

28

Kelsey Forde CIH CSP CHMM
- klforde@parvaticorp.com
- (505) 967-8917

Timothy Stirrup CSP CHMM REM
- tsstirrup@parvaticorp.com
- (505) 980-3743

29