AGENDA ITEM #11

BRIEF INTRODUCTION TO AI TECHNOLOGIES

OFFICE OF CYBER DEFENSE COORDINATION

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

What is it?

Artificial intelligence is the simulation of human intelligence processes by machines, especially computer systems.

ARTIFICIAL INTELLIGENCE / AI

How does it work?

AI systems work by ingesting large amounts of labeled training data, analyzing the data for correlations and patterns, and using these patterns to make predictions about future states.

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

Current Types of AI

- Machine Learning
- LLMs – Large Language Models
- Generative AIs
- GPTs

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

Current Types of AI – **Machine Learning**

- Been around since 1959.
- What all modern AIs are fundamentally based on
- Mathematical models and algorithms are fed specific data sets and trained to identify patterns within each set, which they can use to predict when patterns and actions are likely to happen given similar input.

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

Current Types of AI – **LLMs – Large Language Models**

- A newer type of machine learning model where it is trained using “unsupervised learning”, which means the algorithm is given a data set, but isn’t programmed on what to do with it.
- Through this process, an LLM learns how to determine the relationship between words (or images, or whatever) and the concepts behind them.
- Can generate a lot of insights not immediately obvious to the developers of the model.

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

Current Types of AI – **Generative AI**

- A LLM that is capable of taking the connections it has made and using it for generating content such as text, video or audio based on input provided and pattern matching predictions.
- These are what started making the news as AI, under the term “Deep Fakes”

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

Current Types of AI – **GPT**

- Generative Pre-Trained Transformer
- A specific form of Generative AI that comes with pre-built rules around its generative content creation.
- Allows for very directed, very “intelligent” content creation.
- ChatGPT, StableDiffusion, Midjourney, and the like all fall into this category.

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

What's the good?

- **Details aren't missed!** AI has proven to be very effective at diagnosing certain cancers or other medical issues, given a complete enough data set on the patient.
- **Reduced time** for data-heavy tasks.
- **Increased productivity** in tasks that can be automated.
  - This includes business processes; warehouse automation is a big industry making use of AI
- **Delivers consistent results,** with consistent input.

ARTIFICIAL INTELLIGENCE / AI

What's the bad?

- **Expensive,** in resources and hardware.
- **Requires deep technical expertise** to train effectively.
- **Limited supply of qualified workers to build AI tools.**
- **Reflects the biases** of its training data, and prompts, at scale.
- **Lack of ability to generalize** from one task to another.
- **LLMs are trained to give you an answer, no matter what.** If it doesn't know; it will just make something up!

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

Current worries and concerns

- Employment Impact
- Inadvertent information leaks
- Ethical implications of content scraping
- Unvalidated content creation
- Misuse

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

How is it being used: Healthcare

- IBM Watson is becoming a very strong reference for diagnoses.
- AI Chatbots are used for scheduling appointments
- Predict disease progress at scale (flu, covid infections, etc)

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

How is it being used: Business

- CRM
- Customer Service chatbots
- Document summaries and quick content creation

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

How is it being used: Finance

- High speed brokerage trading
- Personal financial advice tools
- Fraud detection and analytics

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

How is it being used: Law

- Automate legal research
- Generate case summaries
- Document digitalization and indexing

ARTIFICIAL INTELLIGENCE / AI

How is it being used: Entertainment and Media

- Content recommendation algorithms
- Automated content creation
- Filling in backgrounds in movies
- Creating voice/video representations of dead actors

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

How is it being used: IT and Cybersecurity

- Automating log analysis and event detection
- Code generation assistance
- Automated assessments and evaluations

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

How can we use it well?

- Keep in mind what data we feed into non-local tools
- Use it to assist in research and summaries, but validate
- Keep in mind implications of assisted content creation

AGENDA ITEM #11

ARTIFICIAL INTELLIGENCE / AI

Further research

- Example use cases in government: https://ai.gov/ai-use-cases/
  - Federal Government is creating a thorough inventory of their use cases
- National AI Advisory Committee reports: https://ai.gov/naiac/
  - "Rationales, Mechanisms, and Challenges to Regulating AI" discusses challenges and implications
- Known security considerations: OWASP Top 10 for LLM

https://owasp.org/www-project-top-10-for-large-language-model-applications/assets/PDF/OWASP-Top-10-for-LLMs-2023-v1_1.pdf

AGENDA ITEM #11

THANK YOU!

ANY QUESTIONS?