GMIS NEW JERSEY
An Association Of Government IT Leaders

LEAGUE OF MUNICIPALITIES INTRODUCTION TO ARTIFICIAL INTELLIGENCE TECHNOLOGY

BY: JOHN J TUGMAN, COUNTY OF MORRIS CIO, GMIS BOARD MEMBER

GMIS NEW JERSEY
An Association Of Government IT Leaders

TRADITIONAL & GENERATIVE AI

- **Traditional AI - focuses on analyzing historical data and making future numeric predictions.**

- **Generative AI - allows computers to produce brand-new outputs that are often indistinguishable from human-generated content.**

Tech bubble is *bigger* than the 1990s dot-com bubble.

BY: JOHN J TUGMAN, COUNTY OF MORRIS CIO, GMIS BOARD MEMBER

Generative AI's evolution

For an advanced technology that's considered relatively new, generative AI is deep-rooted in history and innovation.

1932
Georges Artsrouni invents a machine he reportedly called the "**mechanical brain**" to translate between languages on a mechanical computer encoded onto punch cards.

1966
MIT professor Joseph Weizenbaum creates the first chatbot, **Eliza**, which simulates conversations with a psychotherapist.

1980
Michael Toy and Glenn Wichman develop the Unix-based game *Rogue*, which uses procedural content generation to dynamically generate new game levels.

1986
Michael Irwin Jordan lays the foundation for the modern use of recurrent neural networks (RNNs) with the publication of "Serial order: a parallel distributed processing approach."

2000
University of Montreal researchers publish "A Neural Probabilistic Language Model," which suggests a method to model language using feed-forward neural networks.

2011
Apple releases **Siri**, a voice-powered personal assistant that can generate responses and take actions in response to voice requests.

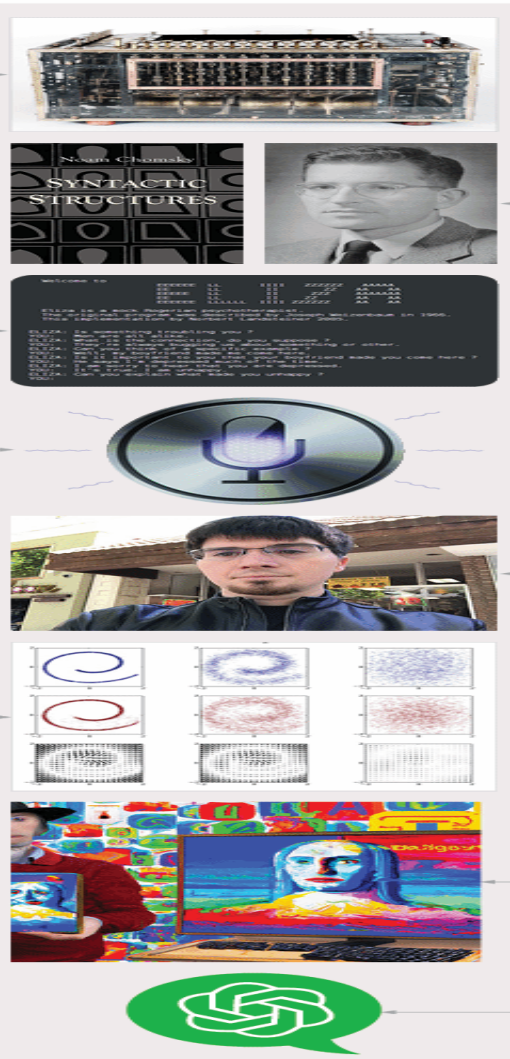
2013
Google researcher Tomas Mikolov and colleagues introduce word2vec to identify semantic relationships between words automatically.

2015
Stanford researchers publish work on diffusion models in the paper "**Deep Unsupervised Learning using Nonequilibrium Thermodynamics**." The technique provides a way to reverse-engineer the process of adding noise to a final image.

2018
Google researchers implement transformers into BERT, which is trained on more than 3.3 billion words and can automatically learn the relationship between words in sentences, paragraphs and even books to predict the meaning of text. It has 110 million parameters.

Google DeepMind researchers develop AlphaFold for predicting protein structures, laying the foundation for generative AI applications in medical research, drug development and chemistry.

OpenAI releases GPT (Generative Pre-trained Transformer). Trained on about 40 gigabytes of data and consisting of 117 million parameters, GPT paves the way for subsequent LLMs in content generation, chatbots and language translation.



1957
Linguist **Noam Chomsky** publishes ***Syntactic Structures***, which describes grammatical rules for parsing and generating natural language sentences.

1968
Computer science professor Terry Winograd creates SHRDLU, the first multimodal AI that can manipulate and reason out a world of blocks according to instructions from a user.

1985
Computer scientist and philosopher Judea Pearl introduces Bayesian networks causal analysis, which provides statistical techniques for representing uncertainty that leads to methods for generating content in a specific style, tone or length.

1989
Yann LeCun, Yoshua Bengio and Patrick Haffner demonstrate how convolutional neural networks (CNNs) can be used to recognize images.

2006
Data scientist Fei-Fei Li sets up the ImageNet database, which provides the foundation for visual object recognition.

2012
Alex Krizhevsky designs the AlexNet CNN architecture, pioneering a new way of automatically training neural networks that take advantage of recent GPU advances.

2014
Research scientist **Ian Goodfellow** develops generative adversarial networks (GANs), which pit two neural networks against each other to generate increasingly realistic content.

Diederik Kingma and Max Welling introduce variational autoencoders to generate images, videos and text.

2017
Google researchers develop the concept of transformers in the seminal paper "Attention is all you need," inspiring subsequent research into tools that could automatically parse unlabeled text into large language models (LLMs).

2021
OpenAI introduces **Dall-E**, which can generate images from text prompts. The name is a combination of WALL-E, the name of a fictional robot, and the artist Salvador Dali.

2022
Researchers from Runway Research, Stability AI and CompVis LMU release Stable Diffusion as open source code that can automatically generate image content from a text prompt.

OpenAI releases **ChatGPT** in November to provide a chat-based interface to its GPT 3.5 LLM. It attracts over 100 million users within two months, representing the fastest ever consumer adoption of a service.

2023
Getty Images and a group of artists separately sue several companies that implemented Stable Diffusion for copyright infringement.

Microsoft integrates a version of ChatGPT into its Bing search engine. Google quickly follows with plans to release the Bard chat service based on its Lamda engine. And the controversy over detecting AI-generated content heats up.

DESIGN: LINDA KOURY; CHOMSKY: BETTMANN/GETTY IMAGES; ELIZA: WIKIMEDIA COMMONS; IAN GOODFELLOW: WIKIMEDIA COMMONS; DALL-E: SHAUN SUTNER

TERMS

- **Hallucinations: When an AI model presents inaccurate information. The production of confidently stated but erroneous or false content also known as Confabulation.**
- **LLM's (Large Language Models) - A large language model is AI trained on a lot of text data. It can understand, generate, and predict new content.**
- **Generative AI: A type of AI that can think, learn, perform intellectual tasks that humans do, and execute tasks it's not trained to do.**
- **GPT: "Generative pre-trained transformer," refers to both a specific model and a family of progressively more sophisticated artificial intelligence (AI) models.**
- **ChatGPT (Chat Generative Pre-trained Transformer) — is a large language model (LLM)-powered chatbot developed by OpenAI. It enables users to interact with it using human-like conversation.**
- **Copilot: Microsoft 365's AI assistant feature that builds on OpenAI's GPT-4 large language models (LLMs).**

TERMS (CONT.)

- **Neural Network: a neural network is a subset of machine learning. It mimics the way the human brain's neurons signal to one another to solve a problem or deliver an answer.**
- **Transformer: is a deep learning model type for natural language processing. It can process the context of words in a sentence and produce an output based on a sequence of data throughout a conversation.**
- **CDAO (Chief Digital and Artificial Intelligence Office): h as five primary functions: · Lead and oversee strategy development and policy formulation for data, analytics, and AI. Primarily responsible for generating value through the organization's data and analytics assets and ecosystem.**
- **AI Doomerism: belief in the likelihood of AGI causing the downfall of humanity resulting from a superhuman intelligence.**
- **Accelerationists (e/accs) say we need to build AI as quickly as possible, "Even if it risks wiping out humanity, that is a good thing because it means that intelligence will advance in the universe."**
- **Convolutional Neural Network (CNN): A type of artificial neural network used primarily for image recognition and processing, due to its ability to recognize patterns in images.**

TERMS (CONT.)

- **DALL-E: an AI system that uses machine learning to create images and art from a user's description. Its most recent version, DALL-E 2, has higher-resolution images.**
- **Deepfake: an AI-generated image, audio, or video depicting fake events. It uses powerful machine learning and AI to manipulate or create deceiving content.**
- **Generative Adversarial Network (GAN): a type of machine learning that comprises two neural networks (a generator and a discriminator) competing. The generator creates an output based on an input, while the discriminator identifies whether the output is real or not.**
- **Anthropomorphize: attribute human characteristics or behavior to (a god, animal, or object).**
- **Retrieval-augmented Generation(RAG): draws upon external data sources to address two shortcomings of large language models, out-of-date training sets and limited context windows.**

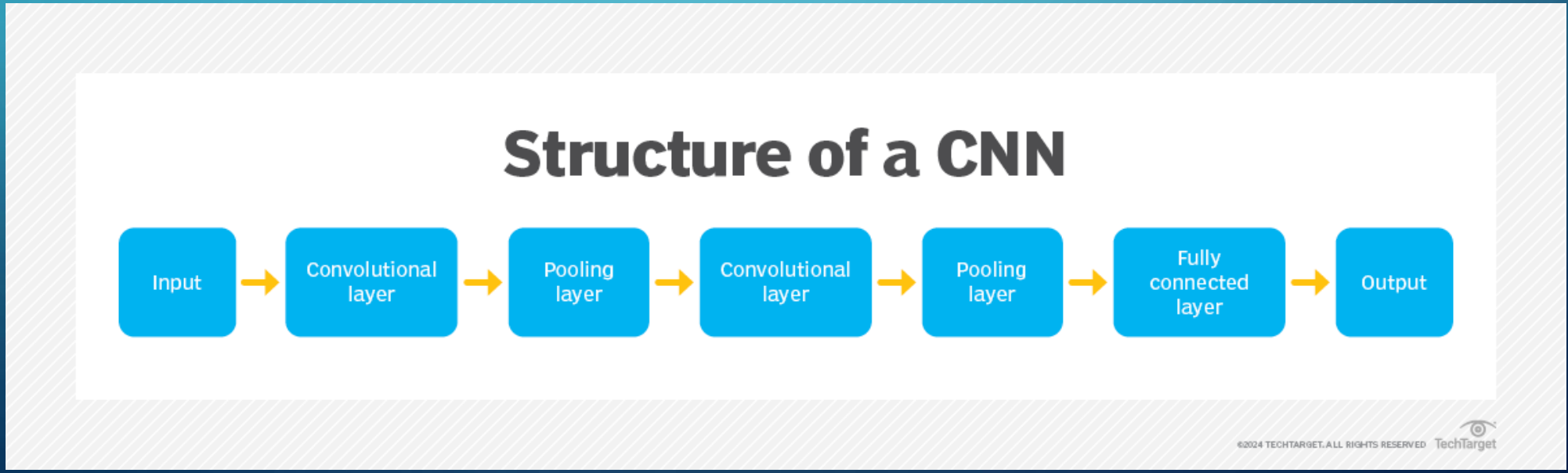
GENERATIVE AI – "HOW IT WORKS"

Convolutional Neural Network (CNN): A convolutional neural network (CNN) is a type of artificial neural network used primarily for image recognition and processing, due to its ability to recognize patterns in images. A CNN is a powerful tool but requires millions of labelled data points for training.

CNN: Consists of three groups:

1. **Convolutional layers**
2. **Pooling layers**
3. **Fully connected layers.**

As data passes the groups/layers the complexity of the CNN increases, which lets the CNN successively identify larger portions of an image and more abstract features.



GMIS NEW JERSEY
An Association Of Government IT Leaders

GENERATIVE AI – HARDWARE

- **Core Components/AI Hardware:**
  - Central Processing Unit (CPU)
  - Graphics Processing Unit (GPU)
  - Tensor Processing Unit (TPU):is an AI accelerator application-specific integrated circuit (ASIC) developed by Google for neural network machine learning. More efficient than adding 100's of GPU's
    - Field-Programmable Gate Arrays (FPGAs): Programable hardware logic.
  - Neural Network Processors (NNPs):

GMIS NEW JERSEY
An Association Of Government IT Leaders

AI HARDWARE (CONT'D)

Products Examples

- **Microsoft - Copilot+ PCs**
- **HP Omni Book X AI PCs**
- **Lenovo - ThinkPad's T14's & X 1's and Think Books Gen 4's, 5's & 6's**

GMIS NEW JERSEY
An Association Of Government IT Leaders

AI PROS AND CON'S

GMIS NEW JERSEY
An Association Of Government IT Leaders

PROS

- Reduced time to market New Medications: develop more targeted medicines, driving progress toward precision medicine.
- AI is estimated to save 5% to 10% in healthcare spending, which is equivalent to annual savings of US$200 billion to US$360 billion (Healthcare Dive, 2023).Oct 17, 2023
- **View generative AI as a tool to augment human capabilities, not as a replacement.** "When used strategically and responsibly, generative AI can transform efficiency, creativity, and decision-making, driving innovation and competitive advantage."
- Provide flexible and efficient ways to automate tasks, with support for significant customization.

GMIS NEW JERSEY
An Association Of Government IT Leaders

CON'S

- Risks to data privacy resulting from the generation of vast amounts of sensitive patient data.
- Generative AI enables nearly anyone to become a sophisticated cybercriminal in a matter of seconds, providing tips on how to get started, and the exact elements needed to execute a successful attack increasing the number of people creating attacks.
- Bias and fairness concerns in training data that may lead to unequal treatment, misdiagnosis, or underdiagnosis of certain demographic groups.
- Resistance to adoption by healthcare professionals and the public driven by a lack of trust in AI-generated recommendations.

"The opacity of AI has created a significant trust gap that only transparency can fully bridge."

"A chisel in the hands of a trained professional can create amazing things; a chisel in the hands of an amateur can be a lost opportunity."

- Ethical concerns arising from AI-generated decisions that may conflict with patient or family preferences.
- Can't incorporate real-time updates or data after their last training cutoff.
- "Most companies are simply playing with the novelty of AI ."
  - **ROI** missing in the equation.

RISKS

- The output generated by ChatGPT, and other large language model (LLM) tools are prone to several risks, (Reference: Campus Technology, The Dark Side of ChatGPT: 6 Generative AI Risks to Watch, Rhea Kelly 06/02/23):
  - Fabricated and inaccurate answers.
  - Data privacy and confidentiality: Information entered into ChatGPT may become a part of its training dataset.
  - Model and output bias: "Complete elimination of bias is likely impossible," but legal compliance needs to be a primary law governing AI biases.
  - Intellectual property and copyright risks. Because ChatGPT is trained on internet data — including copyrighted material — its outputs "have the potential to violate copyright or IP protections,
  - ChatGPT does not cite sources for the text it generates.

GMIS NEW JERSEY
An Association Of Government IT Leaders

RISKS (CONT'D)

- Cyber fraud risks. "Bad actors misusing ChatGPT to generate false information at scale"
  - Offering fake reviews to influence consumer purchasing decisions.
  - Prompt injection: Hacking technique used to trick the model into performing tasks that it wasn't intended for such as writing malware codes or developing phishing sites that resemble well-known sites.
- Consumer protection risks. Organizations must disclose clearly and conspicuously that a consumer is communicating with a bot.

GENERATIVE AI & DATA MANAGEMENT

- AI can provide government agencies the power to:
  - Identify
  - Curate
  - Activate high-value data to dramatically increase speed-to-mission value.

- https://openai.com/index/hello-gpt-4o/ AI Voice Response Example

GENERATIVE AI – DATA

- Success of AI projects: Fundamentally relies on mass quantities of accessible, reliable **Data**.
  - AI, ML, and analytics output are meaningful **only if the data they operate on is valid and observable** across the whole lifecycle .

- Challenges to scaling AI: Cost, lack of talent, trust and ethics.
- Data quality and availability are the biggest hurdles.
  - 72% of technology executives surveyed in a recent MIT study say that should their companies fail to achieve their AI goals data issues are more likely than not to be the reason.
  - 61% of respondents in an IBM survey said their data is not ready for AI.

GENERATIVE AI & DATA(CONT'D)

Focus on two imperatives:
1. integration
2. data.

- A PWC Study:
  - Top tech related challenge for AI:
    - Identifying the data
    - Collecting/aggregating data from across the organization
    - Ensuring the data's completeness and accuracy.

  Followed closely by making sure all data in AI systems meet **regulatory requirements for privacy and data protection** and integrating AI and analytics systems to gain business insights.

  "With technology tools that help you overcome your data challenges, you can achieve much faster (and much more cost effective) operationalizing of AI."

- **How GPT models are trained:**
  - **Unsupervised training:** Model ingests massive amounts of unlabeled data from varied sources like Wikipedia articles, digital books, and online discussions.
  - **Supervised training:** After the unsupervised phase is complete, GPT models are refined using supervised training. In supervised training, humans train the model using tailored, labeled prompts and responses with the goal of teaching the model which responses humans will likely want and which ones are harmful or inaccurate.

Generative AI & Data(cont'd)

- Domain-specific LLMs: LEGAL-BERT for law, BloombergGPT for finance, and Google Research's Med-PaLM for medicine.
  - LLMs tend to have a limit, generally between 4k and 32k tokens per prompt, which limits how much an LLM can learn on the fly.
- Natural Language Processing (NLP) is one of the hottest areas of artificial intelligence (AI) thanks to applications like text generators that compose coherent essays, and text-to-image programs that produce photorealistic images of anything you can describe.

"AI is fundamentally a power-enhancing technology. We need to ensure that it distributes power and doesn't further concentrate it." Reference: CYBERSCOOP, MAY 28, 2024 "**How AI will change democracy" BY:** BRUCE SCHNEIER

AI O&M

- **AI/ML models are not static:**
  - Require ongoing monitoring and maintenance to ensure performance and reliability.
  - Monitoring for concept drift, model decay, and performance degradation is essential.
  - Regular updates and retraining may be necessary to adapt models to evolving data patterns or changes in the operational environment.
- **Organizations must establish processes to manage version control, model updates, and performance tracking.**
  - Most organizations handle these processes manually. They create manual workflows around retraining data, use new datasets, identify boundary conditions or fringe predictions that don't match the norm, and then make the best guess as to the right time to retrain the model. Clearly, this is an imprecise science that can lead to subpar outcomes.

AI O&M (DATA BACKUP)

- As more organizations integrate AI into their operations, the need to protect an increasing amount of data assets becomes a crucial concern.
  - The current gap in backup and recovery for AI-generated data is a major concern. It poses significant business risks and compliance exposures. Organizations that fail to protect their AI-generated data risk losing out on valuable insights, intellectual property, and competitive advantage.
    - On average only 40 to 50% of orgs perform regular AI data backups. Although 90% of orgs claim AI/ML is important
    - Among organizations that feel AI/ML will improve their ability to recover from ransomware attacks, more than two-thirds feel its automation will improve their overall cybersecurity-recovery RPO and RTO, among others benefits.

GMIS NEW JERSEY
An Association Of Government IT Leaders

AI IMPLEMENTATION CONSIDERATIONS

- Primary **considerations** when implementing AI technologies:
  - Guardrails
  - Ensuring they are secure and protect privacy.
  - Ensuring they are accessible for everyone.
  - Ensuring the tools are responsible, which goes beyond Ethics and Legality; it includes Transparency on how these tools are used and where the data that powers them comes from.
  - How can we derive practical value from AI? What is the most cost-effective way to operationalize it? What about data privacy?

- AI implementations must ensure that AI is applied:
  - Ethically
  - Securely
  - Transparently, in a human-centric manner across all sectors.

BY: JOHN J TUGMAN, COUNTY OF MORRIS CIO, GMIS BOARD MEMBER

GMIS NEW JERSEY
An Association Of Government IT Leaders

10 THINGS TO WATCH OUT FOR WITH OPEN-SOURCE GEN AI

1. **Weird new license terms:** Open-source Gen AI, isn't just code. It's also the training data, model weights, and fine tuning.

2. **Skills shortages:** Self-Explanatory

3. **Jailbreaking:** LLMs are notoriously susceptible to jailbreaking, where a user gives it a clever prompt that tricks it into violating its guidelines and, say, generating malware.

4. **Training data risks:** Information Integrity (Vetted?); Data Privacy (De-anonymization of biometric/health data); Dangerous or Violent recommendation training.

5. **New areas of exposure:** Since a gen AI project is more than just the code, there are more areas of potential exposure. An LLM can be attacked by bad actors on several fronts. They could infiltrate the development team on a poorly governed project and add malicious code to the software itself. But they could also poison the training data, fine tuning, or the weights.

10 THINGS TO WATCH OUT FOR WITH OPEN-SOURCE GEN AI (CONT'D)

6. Missing guardrails: Some open-source groups might have a philosophical objection to having guardrails on their models, or they may believe that a model will perform better without any restrictions.

7. Lack of standards: ISO/IEC 42001 standard for AI management systems, released in December last year.

8. Lack of transparency: No information about the background of the model and how it was developed.

9. Lineage issues: A foundation model uses a problematic training data set and from it, someone creates a new model, so it'll inherit these problems. In fact, those problems may go several levels back and won't be visible in the code of the final model.

10. The new shadow IT: Gen AI projects sometimes fall outside the standard software development processes.

- They might come out of data science teams, or skunkworks.
- Developers might download the models to play with and end up getting more widely used.
- Users themselves might follow online tutorials and set up their own gen AI, bypassing IT altogether.

GMIS NEW JERSEY
An Association Of Government IT Leaders

AI TRAINING RESOURCES

- **Amazon: "AI Ready," a new commitment designed to provide free AI skills training to 2 million people globally by 2025.**
- **Grow with Google program.**


WHEN?
WHO?
WHAT?
WHEN?
WHERE?
WHERE?
When?
WHEN?
Where?
When?
ANY
WHERE?
What?
QUESTIONS?
What?
WHEN?
WHAT?
WHAT?
HOW?
Where?
WHAT?
When?
WHERE?
When?
HOW?
HOW?