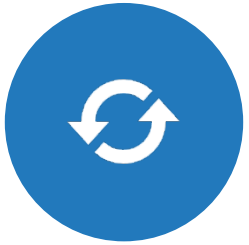
Generative AI 101 – Cyber Hawaii

Understanding the benefits, opportunities and challenges of using GenAI

Matt Leger, MPA
Sr. Research Manager
IDC Worldwide Education and EdTech Digital Strategies
July 2025

Today's Agenda



Introductions

- Introductions
- Goals for Today

GenAI 101

- Understanding different types of AI
- GenAI adoption in education

GenAI Safety

- Consideration for safe and ethical use of GenAI

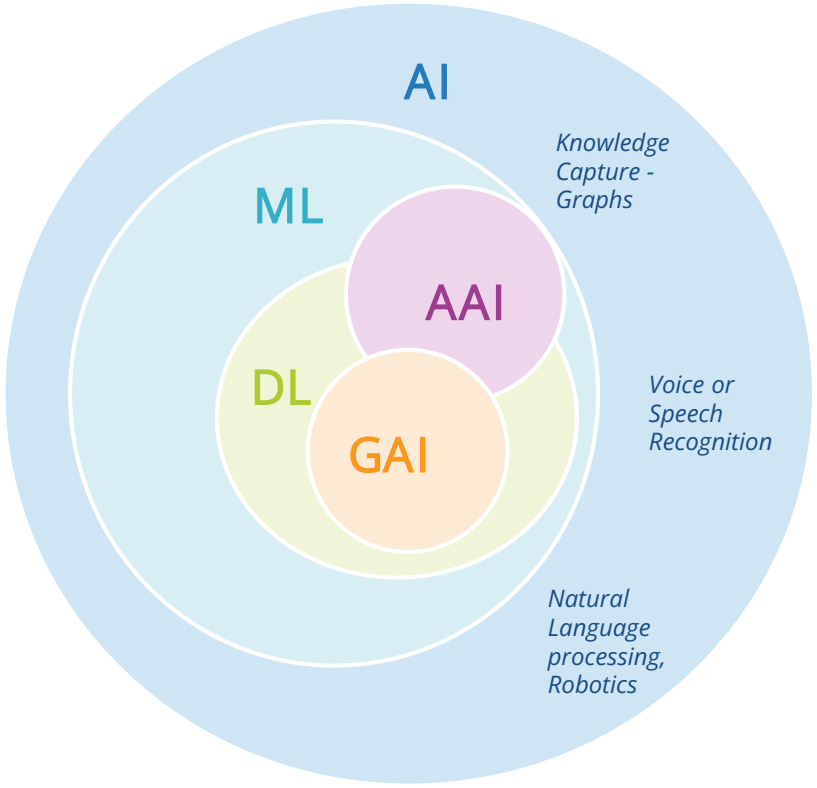
Open Discussion

- Q&A

Foundations of GenAI

GenAI 101

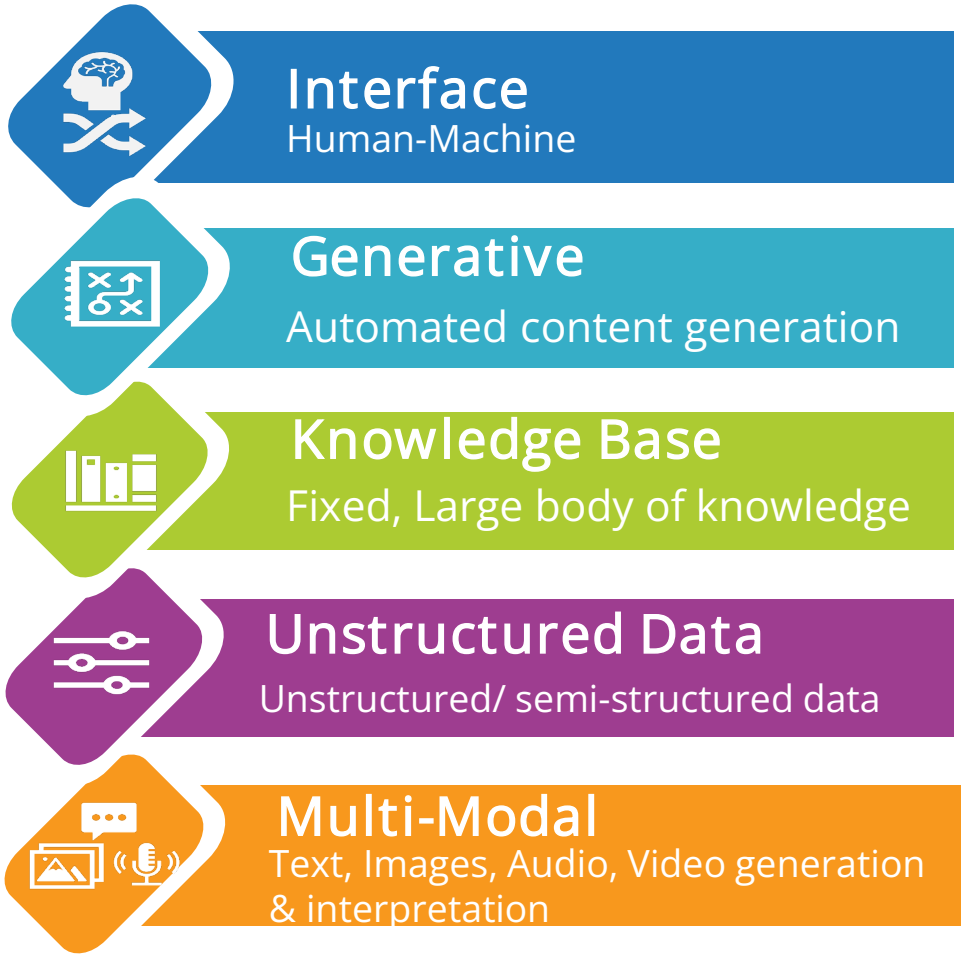
GenAI is a subset of different types of AI



- AI - Artificial Intelligence — *Techniques that help computers mimic human behavior.*
- ML - Machine Learning — *Subset of AI techniques that enable computer systems to learn without programming by a human. E.g., Supervised learning, Unsupervised Learning, Reinforcement Learning.*
- DL - Deep Learning — *Subset of ML techniques that makes the computational multilayer neural networks feasible. E.g., CNN, RNN, GAN.*
- GAI -Generative AI — *Subset of DL techniques that enable computers to **create new content** using previously **created content, such as text, audio, video, images and code.***
- AAI - Agentic AI — *Subset of ML and DL techniques that enable computer systems to exhibit agency: set goals, make decisions and take actions through perception, reasoning and action loop.*

Source: Generative AI and Agentic AI Strategies, April 2025

What makes GenAI so different and exciting?



GenAI follows processes based on **human language.**

GenAI **creates content** from vast amounts of existing data.

GenAI is **driven by trained sets of vast collections of data** organized by proximities of usage. These trained sets are called large language models (LLMs)

GenAI **can leverage unstructured data** such as code, images, video, audio, websites, social feeds, IMs, email, contracts, invoices, manuals, expense reports, conference proceedings, etc.
GenAI **can create content in multiple formats** - text, video, audio. It can create presentations, documents, new graphics etc.

Generative AI is designed to find patterns in both structured and unstructured data using natural language

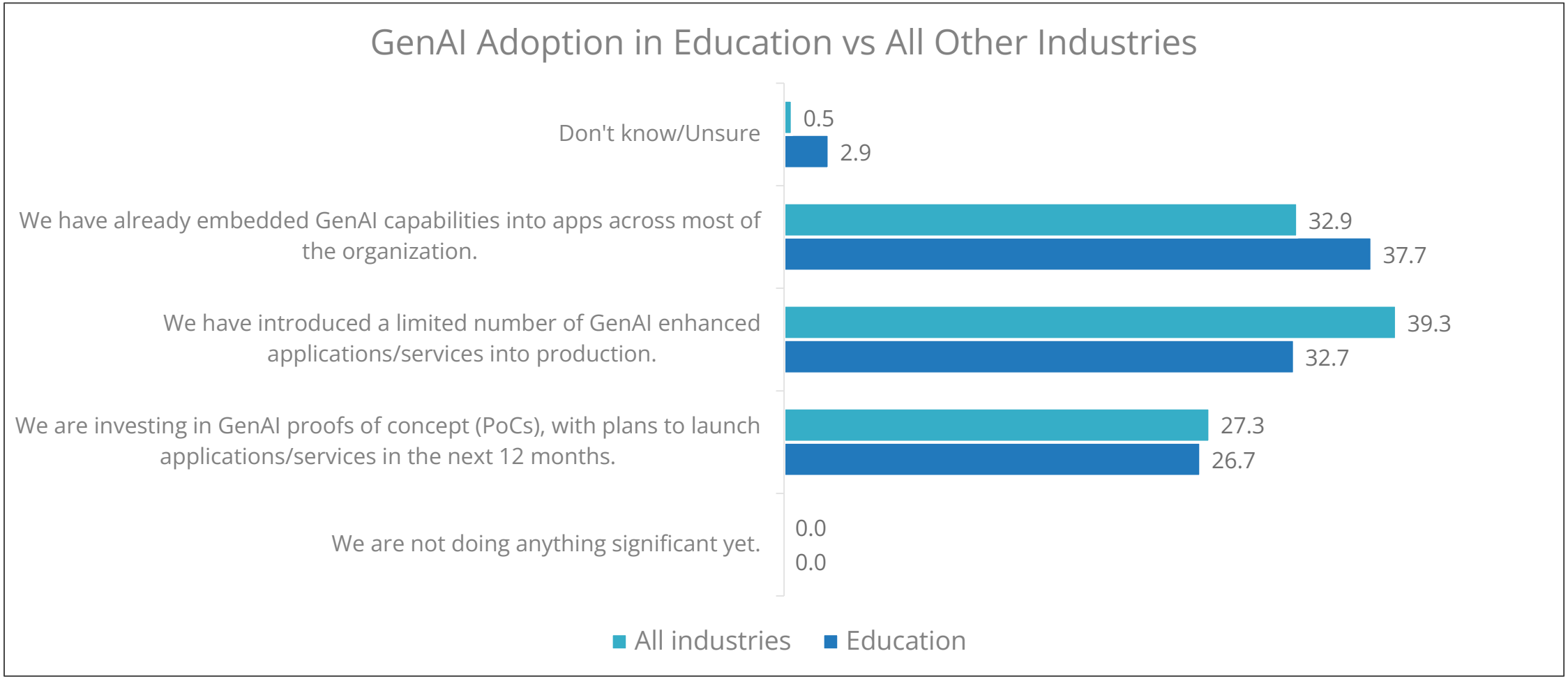
GenAI changed the trajectory of AI adoption. Agentic AI is next…

We are at an inflection point for Autonomous AI Agent Development

AI Co-pilots/AI Assistants	Autonomous AI agents	AI Agent Fleets
AI provides Insights or recommendations *E.g. Order Status Inquiry*	Humans are "architects and orchestrators“ of workflows *E.g., Order Refunds/Exchanges*	Complex problem solving, collaboration across agents *E.g.,* Fully Autonomous Customer Service

Source: Generative AI and Agentic AI Strategies, April 2025

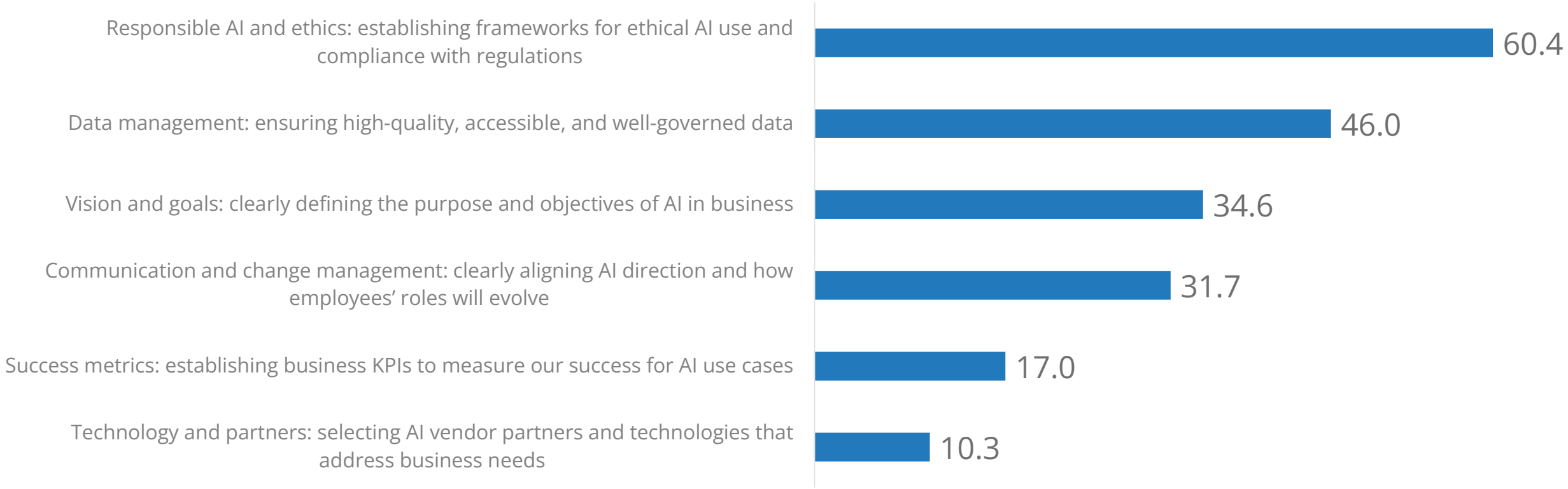
GenAI is no longer optional in K-12 schools. Education is keeping pace with all other industries in adoption.



Future Enterprise Resiliency & Spending Survey Wave 1, IDC, February 2025 (n=38 education, 885 all industries)

As GenAI adoption increases, safe and responsible AI use is priority number one for schools

You indicated that your organization is prioritizing AI strategy in 2025. What is the most important issue that your AI strategy needs to address?



Future Enterprise Resiliency & Spending Survey Wave 1, IDC, February 2025 (n=38 education)

Considerations for Safe and Ethical Use of GenAI

8 Leading Considerations for Safe and Ethical GenAI Use

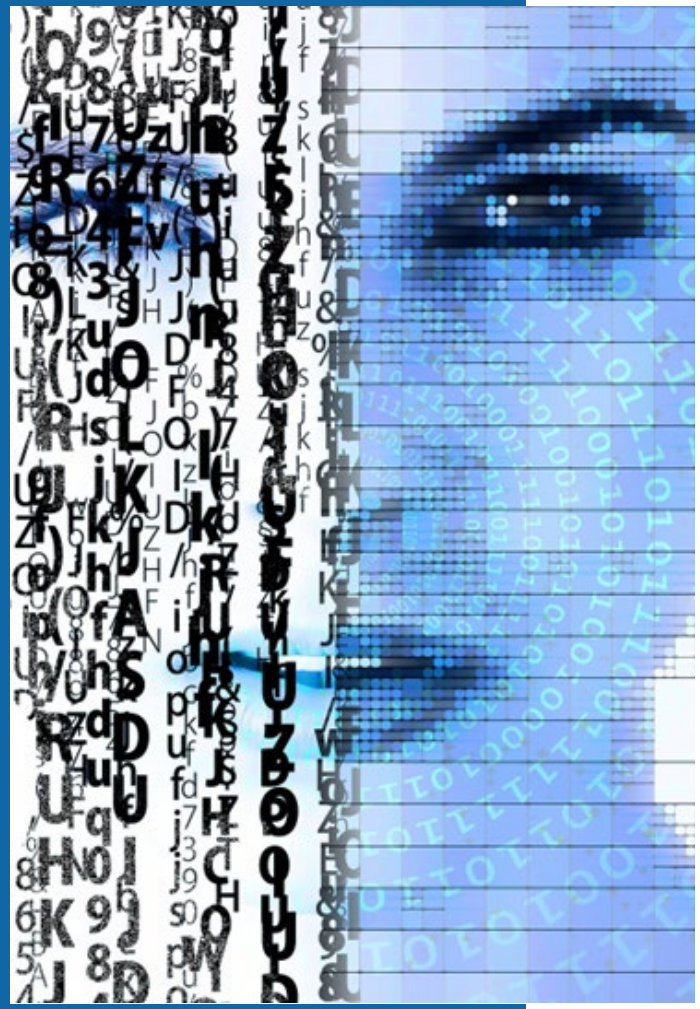
Privacy and Consent

With the explosion of GenAI comes serious questions about privacy, consent, digital identity, and individual rights, particularly for young students.

gt government technology

Where's DeepSeek Banned? The States Blocking Chinese-Made AI

States are increasingly banning DeepSeek AI on government devices, citing cybersecurity and data privacy concerns. Some cybersecurity experts question if the state bans will do enough to protect American data.



Bias and Toxicity in AI Models

Key Issues:

- **Seed Data**: If you enter sensitive or personally identifiable information into the models, that data can become part of the model and visible to others.
- **Input Bias**: Prompts/Inputs asking essentially the same question but with different levels of craftsmanship can generate different responses.
- **Demographic Bias**: AI is a window into the bias of the people who created it.

STATESCOOP

Civil rights experts warn of AI's potential to harm the public

Clarence Okoh, a senior attorney at Georgetown's Center for Privacy and Technology, pointed to a predictive policing program in Pasco County, Florida, which was discontinued last March, that used a generative AI tool to comb through student data and flag at-risk youth, leading to increased surveillance and discipline.

Harmful Content, Misinformation, Fraud, and Scams

FOX NEWS

GenAI, the future of fraud and why you may be an easy target

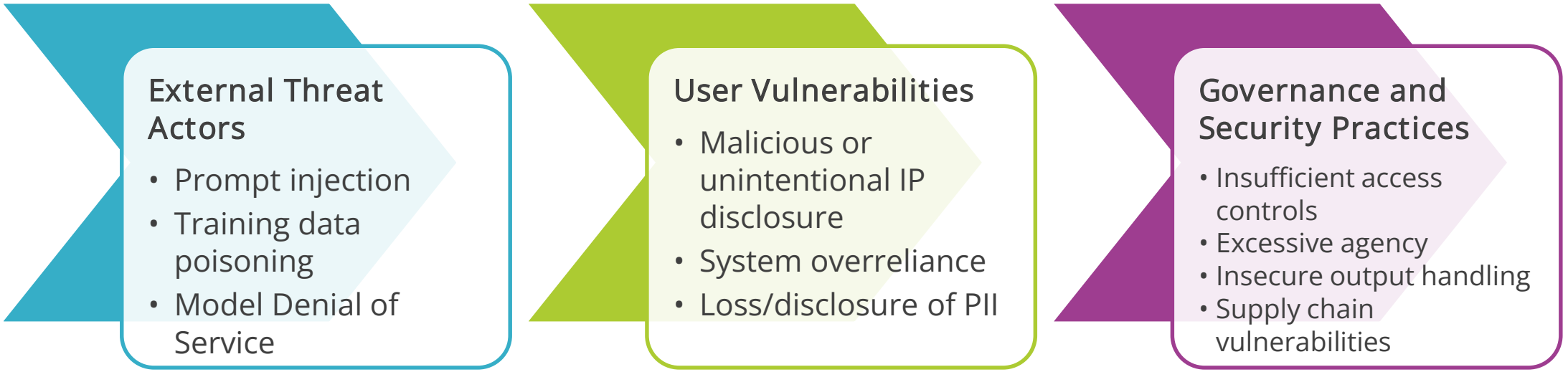
FORTUNE

Job applicants are using deepfake AI to trick recruiters—Here's how hiring managers can spot the next imposter

News4JAX

Jacksonville lawmaker pushes bill aiming to protect victims from deepfakes after former mayor's daughter targeted

Security Vulnerabilities for LLM Applications



The Greatest Threat May Come from Within…

TECH BREW

Workers are using AI faster than their employers can make rules about it

That might cause security and quality concerns, a KPMG report found.

Half of US respondents said they tapped AI at work, despite not knowing whether it's allowed, and 44% said they're "knowingly using it improperly." That includes uploading sensitive information or intellectual property to public AI platforms, which 46% of those in the US admitted to doing.

The survey also pointed to the potential for slipping quality due to AI use. Around three in five (64%) of Americans surveyed "admit to putting less effort into their work, knowing they can rely on AI;" 58% said they don't thoroughly vet outputs; and 57% have made mistakes at work as a result.

Intellectual Property and Copyright Law Violations

How Will Copyright Law and Plagiarism Change in the Age of GenAI?

With the modern Internet, it's easier than ever before to learn from, imitate and even plagiarize other people's work.

Explainability and Transparency

How is GenAI producing the content and output?

NIST's Four Principles of Explainable AI ("XAI)"

01 **Explanation**: Systems deliver accompanying evidence for all outputs

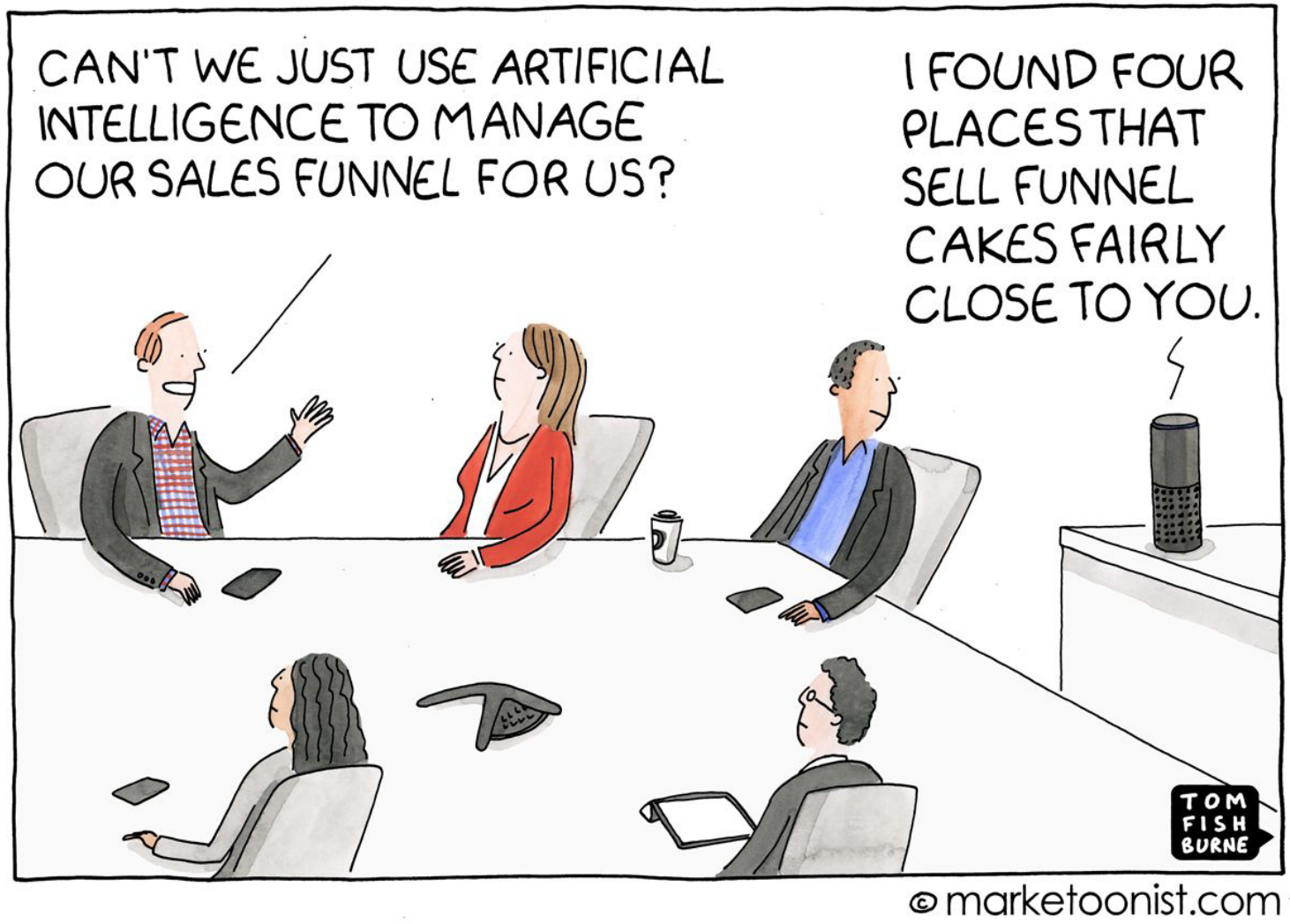
02 **Meaningful**: Systems provide explanations that are understandable to individual users

03 **Explanation Accuracy**: The explanation correctly reflects the system's process for generating the output

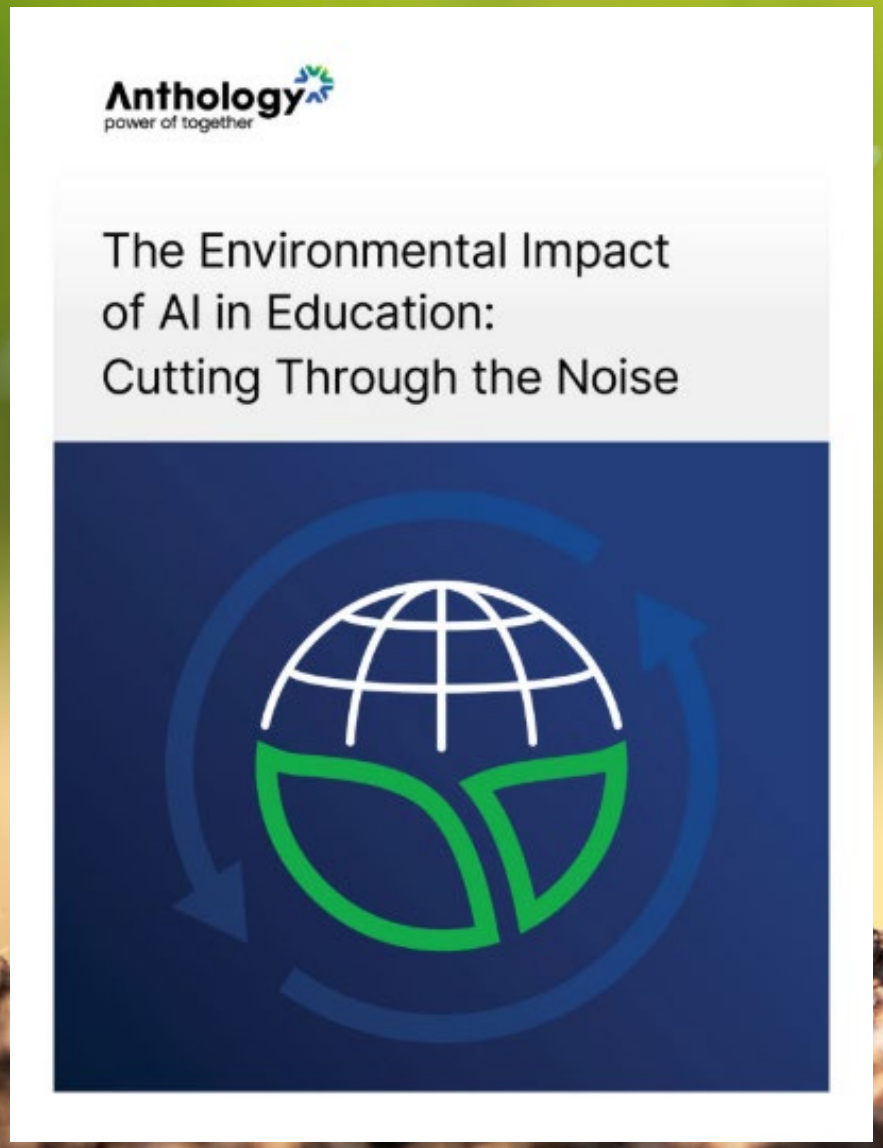
04 **Knowledge Limits**: The system only operates under conditions for which it was designed or when the system reaches a sufficient confidence in its output.

Accuracy and Hallucinations

While some models are improving, GenAI is inherently imprecise



The Environmental Impacts of AI



Human in the loop is key, responsibility falls to the user

Tech/AI companies

- Strong commercial interests
- Varying approaches to AI ethics and responsibility
- Lobbying power

Business and commercial sector

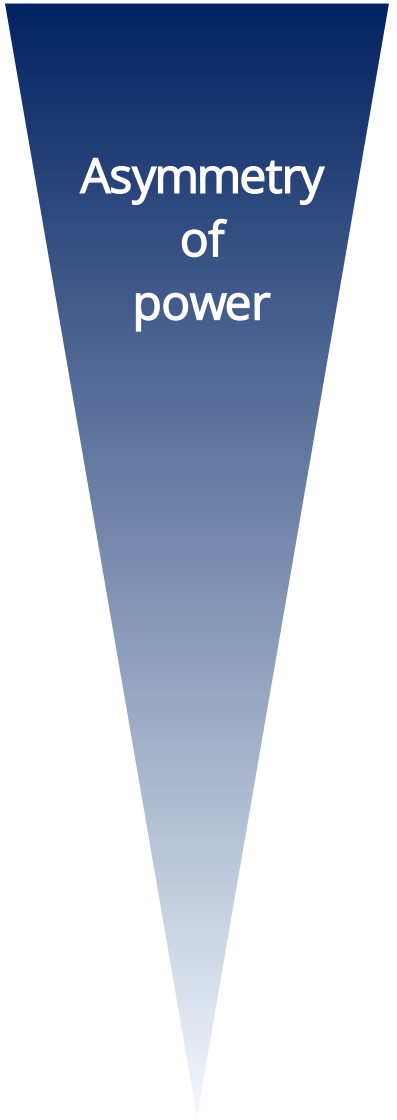
- Predominantly commercial interests
- Efficiency, efficiency, efficiency
- Pushbacks only temporary
- Creative industries might feel more threatened

Governments and public sector

- Expected to act, mitigate negative externalities
- Need more training and access to expertise
- Facing the brunt of public pushback

Individuals: citizens, employees, faculty, students etc…

- Should be at the center of any action
- The right to explain or the right to know
- Need awareness and opt-out option
- The right to license your digital-twin
- Education and critical thinking are key



https://blogs.microsoft.com/on-the-issues/2023/09/07/copilot-copyright-commitment-ai-legal-concerns/
https://nymag.com/intelligencer/article/sam-altman-artificial-intelligence-openai-profile.html
https://corporateeurope.org/en/2023/09/lobbying-power-amazon-google-and-co-continues-grow

The problem with relying on people…

Microsoft | Research

The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers

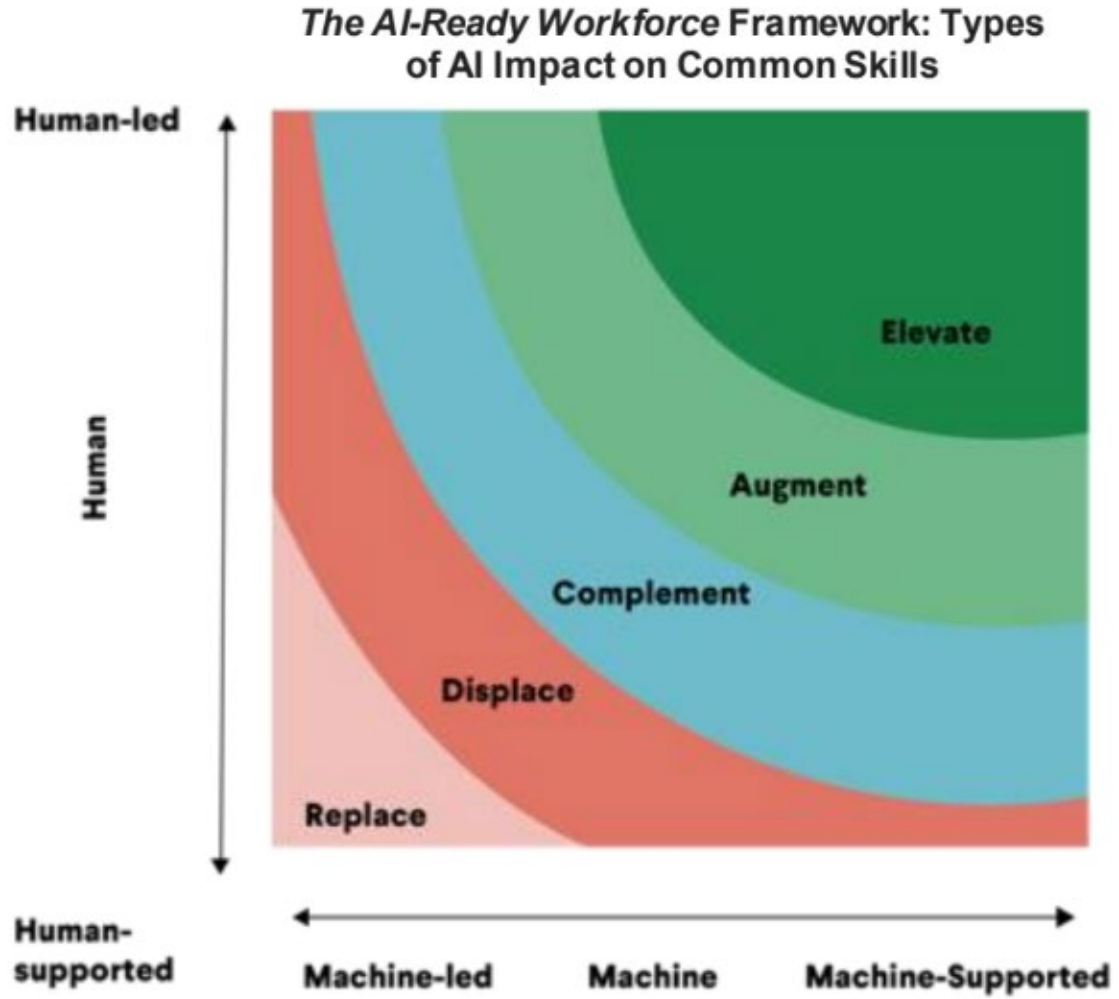
Hao-Ping (Hank) Lee, Advait Sarkar, Lev Tankelevitch, Ian Drosos, Sean Rintel, Richard Banks, Nicholas Wilson

CHI 2025 | April 2025

The rise of Generative AI (GenAI) in knowledge workflows raises questions about its impact on critical thinking skills and practices. We survey 319 knowledge workers to investigate 1) when and how they perceive the enaction of critical thinking when using GenAI, and 2) when and why GenAI affects their effort to do so. Participants shared 936 first-hand examples of using GenAI in work tasks. Quantitatively, when considering both task- and user-specific factors, a user's task-specific self-confidence and confidence in GenAI are predictive of whether critical thinking is enacted and the effort of doing so in GenAI-assisted tasks. Specifically, higher confidence in GenAI is associated with less critical thinking, while higher self-confidence is associated with more critical thinking. Qualitatively, GenAI shifts the nature of critical thinking toward information verification, response integration, and task stewardship. Our insights reveal new design challenges and opportunities for developing GenAI tools for knowledge work.

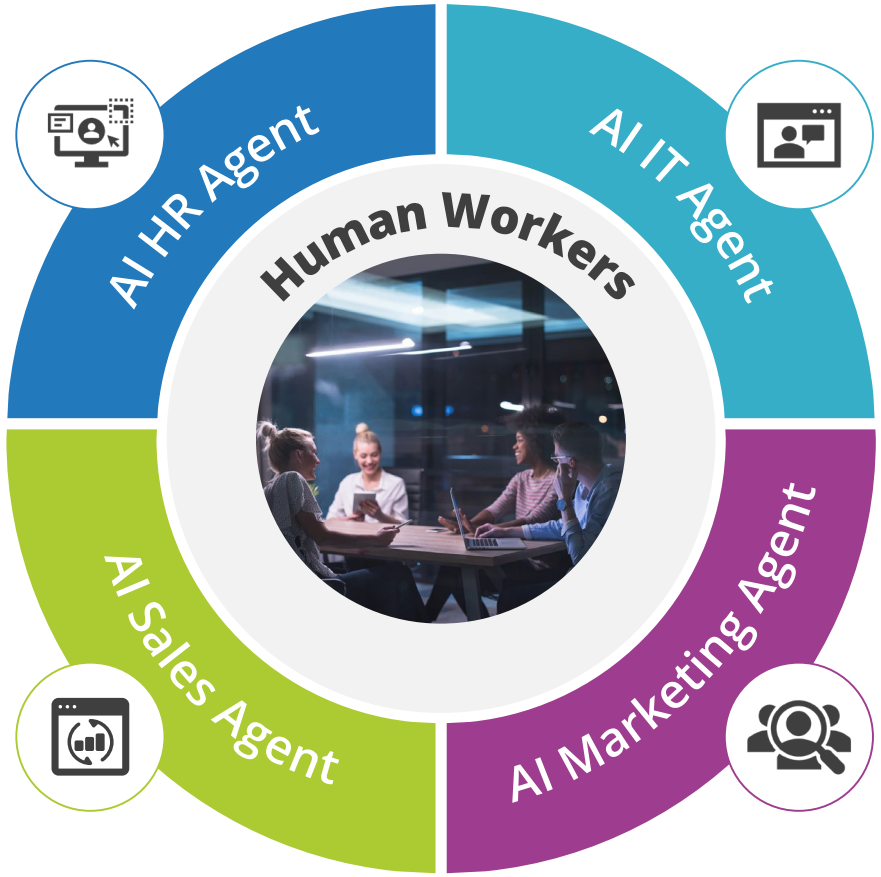
A workforce equipped with AI + human skills will be ready for the future and equipped to use AI ethically and responsibly…

Jobs for the Future



Adoption of agentic AI technologies will inevitably create new job roles and skills requirements

Functional Digital Agents



Networks of AI agents will work with human workers and independently

AI-Agents will take charge of:

- Executing repetitive tasks
- Data analysis and outcomes evaluation
- Acting on behalf of humans to take specific actions
- Generating recommendations for human decisions

Human technical skills will be focused on:

- Initiating requests for agents
- Critical evaluation of AI output
- Learn to orchestrate agentic workflow
- Innovating new products, services

Source: *IDC's Future of Work Research, 2025*

Open Discussion/Q&A

Key AI Terms and Definitions

Types of AI

- **Artificial Intelligence (AI):** AI is the technology that enables computers and machines to perform cognitive functions that simulate human intelligence (thinking, reasoning, remembering).
- **Machine Learning (ML)**: A form of AI that uses algorithms that are trained on data to detect patterns and learn to make predictions and recommendations.
- **Algorithm:** A set of instructions we give to computers that is designed to accomplish a task.
- **Natural Language Processing (NLP)**: A form of AI/ML that gives computers the ability to understand and interpret natural human language by analyzing large volumes of voice and text data.

Key Areas Related to AI

- **AI Governance**: The processes, policies, and structures put in place to ensure the responsible and use of AI systems within an organization or society.
- **AI Ethics**: The study of the moral and societal implications of AI, focusing on issues such as fairness, transparency, accountability, and privacy.
- **Responsible AI**: The practice of designing, developing, and deploying AI systems in a manner that prioritizes ethics, fairness, transparency, and accountability.
- **Human-in-the-loop**: The practice of ensuring that human beings provide a final check against AI-generated outputs/content with the aim of reducing risks from misinformation, copyright infringement, or other issues.

Technical Definitions

- **Structured data:** data that is highly specific and is stored in a predefined format. Examples are forms (e.g. invoicing systems), excel files, data in a preset structure that makes it easier to catalogue and find (e.g. contact lists),databases.
- **Unstructured data:** data that is in many varied types that are stored in their native format. Examples are video, Word documents, audio files, social media and emails. They take up more space and require special expertise to use for analysis and intelligence.
- **Large language models:** Large language models (LLMs) are a type of artificial intelligence (AI) program that can perform a variety of natural language processing (NLP) tasks. They are built on machine learning and trained on large amounts of data.
- **Prompt engineering:** Prompt engineering is the practice of using engineering methodologies to develop prompts that enable users to efficiently use LLMs.