

# Generative AI risks and insurance considerations

Jaymin Kim

Senior Vice President, Emerging Technologies

# Generative AI in the headlines



## Google Cloud is partnering with Mayo Clinic as it tries to expand use of generative A.I. in health care

PUBLISHED WED, JUN 7 2023-8:00 AM EDT | UPDATED WED, JUN 7 2023-10:33 AM EDT



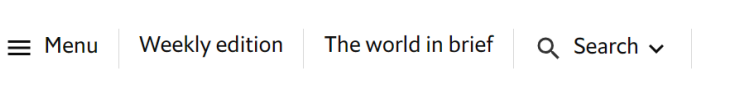
## ChatGPT invented a sexual harassment scandal and named a real law prof as the accused

The AI chatbot can misrepresent key facts with great flourish, even citing a fake Wash



Home > Blog > AI scientists are studying the "emergent" abilities of large language models

## AI scientists are studying the "emergent" abilities of large language models

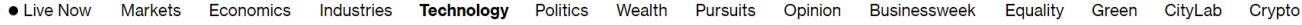


Business | First thing we do, let's bot all the lawyers

## Generative AI could radically alter the practice of law



US Edition Sign In Subscrib



## GPT AI Has Arrived in the Doctor's Office

A new tool from Carbon Health aims to reduce a doctor's workload by generating records and billing codes.



## How Hackers Can Up Their Game Using ChatGPT

Artificial intelligence, by mimicking the writing style of individuals, can make cyberattacks much harder to detect

By Cheryl Winokur Munk  
June 5, 2023 12:00 pm ET



ARTIFICIAL INTELLIGENCE

## The Unpredictable Abilities Emerging From Large AI Models

Large language models like ChatGPT are now big enough that they've started to display startling, unpredictable behaviors.

# Generative AI use cases across industries

Illustrative

### Financial Services

Morgan Stanley

Bloomberg 

- Enable wealth management
- Expense management
- Finance natural language processing




### Healthcare

 Microsoft PaLM 2

 GLASS




- Power clinical tools
- Answer medical questions
- Generate differential diagnoses

### Entertainment




- Produce and edit images
- Generate video script
- Create music

### Retail




- Summarize customer reviews
- Enable research for buyer teams
- Summarize customer database sets

### Real estate and hospitality

- Enable easier search for property listings
- Personalize travel itineraries
- Virtualize staging

### Sports and events

- Personalize fan engagement
- Automate sports commentary
- Communicate pre-/post-game updates in multiple languages

# Artificial intelligence

## Overview

### Artificial intelligence (AI)

The development of computer systems that can simulate human intelligence

#### Narrow AI

Performs specific tasks in a limited domain(s) and cannot perform tasks beyond the limited domain(s)

#### General AI

Performs tasks that humans typically can perform across any domain

#### Generative AI

Employs primarily deep learning models to generate new content across various domains

# Narrow AI

## Overview

### Techniques

**Knowledge-based systems**

- Represent, store, and manipulate domain-specific knowledge
- Use an inference engine to solve problems, provide recommendations, and answer queries

**Machine learning**

- Algorithm-based
- Use data patterns to learn and perform descriptive, predictive, and prescriptive tasks
- Some models use artificial neural networks (ANNs)

**Deep learning**

- Subfield of machine learning
- Differentiated by its use of multi-layered ANNs to model more complex data patterns, larger amounts of data, and more computational resources

### Domains

**Natural language processing**

- **Objective:** Understand, generate, and interact with human language based on human language inputs

**Image synthesis / generative art**

- **Objective:** Generate new images based on textual or visual inputs

**Computer vision**

- **Objective:** Understand and interpret images, simulating human vision capabilities (image recognition, object detection, 3D reconstruction, etc.)

**Robotics**

- **Objective:** Create autonomous machines that can perform tasks, interact with and adapt to environments, and make decisions

# Generative AI vs. “traditional” AI



## Content creation

Create vs. classify



## Learning ability

Existing data vs. new data



## Human-AI interaction

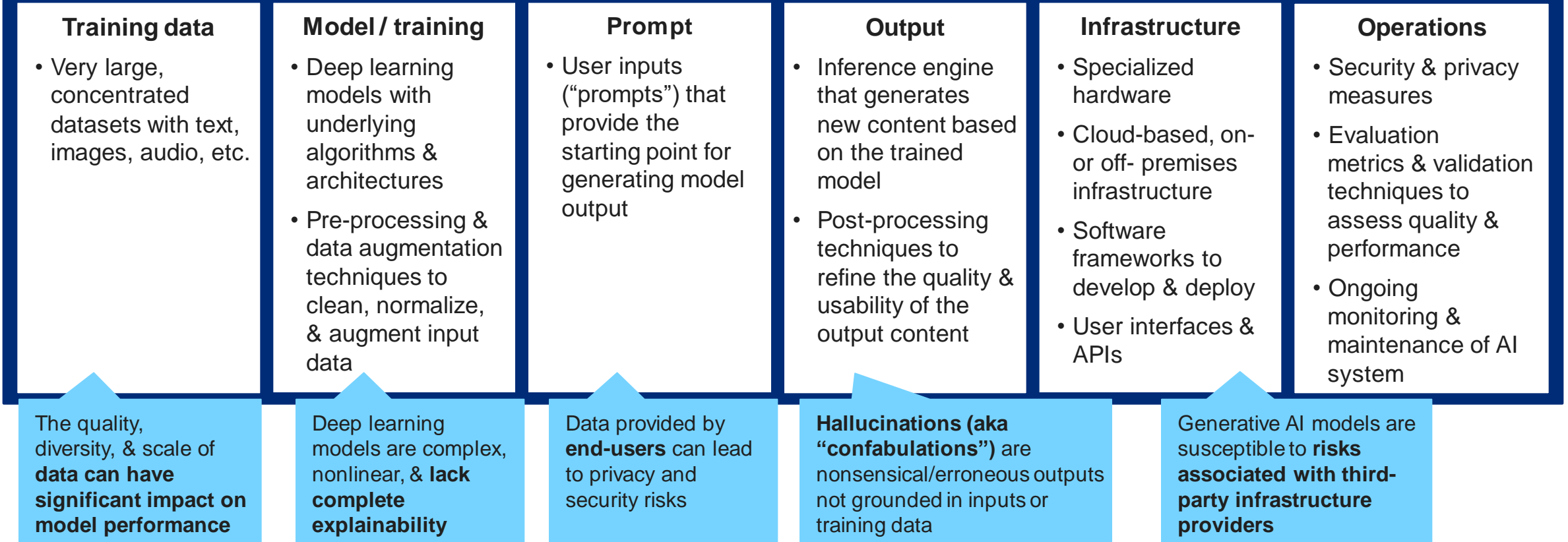
One-way vs. two-way

# Generative AI

Simplified / Illustrative

## Overview

### Components of generative AI



### Regulations

### Contractual Liability

# Generative AI risks framework

## Potential for human error

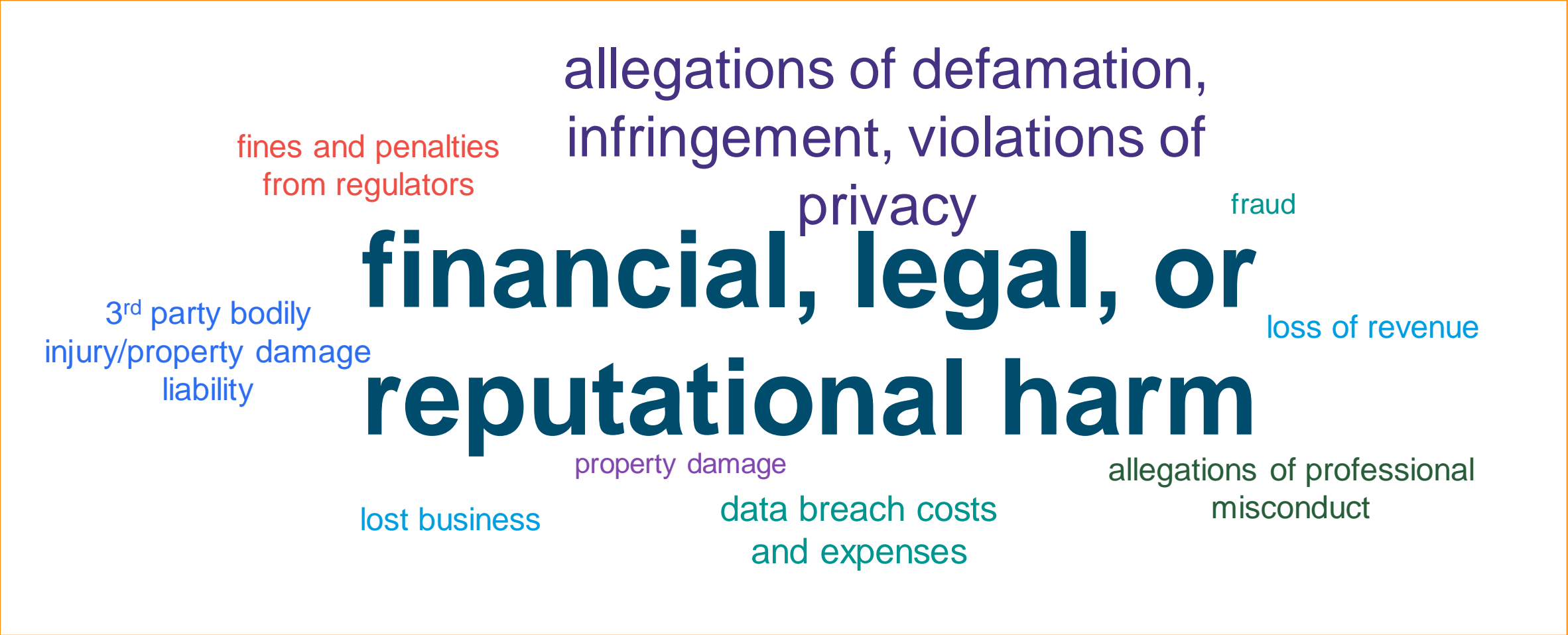
Potential for human error

### Components of generative AI

Training data	Model/ training	Prompt	Output	Infrastructure	Operations
<ul style="list-style-type: none"> <li>• Unauthorized access to, use, or disclosure of confidential personal (PII) training data</li> <li>• Manipulation of data sets by malicious actors</li> <li>• Unauthorized selection/use of copyrighted data/content</li> <li>• Selection of biased, incomplete, inaccurate, or false data for model training</li> </ul>	<ul style="list-style-type: none"> <li>• Implanting malicious code</li> <li>• Selecting model architectures that can generate biased/erroneous outputs</li> <li>• Insufficient cleaning, filtering, or processing of data prior to training</li> <li>• Developing algorithms that lead to biased/erroneous outputs</li> </ul>	<ul style="list-style-type: none"> <li>• Inputting confidential information in environments without adequate security and privacy controls</li> <li>• Inputting prompts to intentionally override built-in system restrictions (“jailbreaking”)</li> </ul>	<ul style="list-style-type: none"> <li>• Accessing, using, or disclosing sensitive, personal output data</li> <li>• Developing/manipulating an inference process that favors specific types of outputs over others</li> <li>• Using post-processing techniques that lead to biased outputs</li> <li>• Misusing model outputs</li> </ul>	<ul style="list-style-type: none"> <li>• Relying on third-party providers without sufficiently assessing counterparty risk</li> <li>• Insufficiently managing/monitoring relevant processes</li> </ul>	<ul style="list-style-type: none"> <li>• Using insufficient security and privacy measures</li> <li>• Using biased/erroneous metrics or techniques to assess model quality &amp; performance</li> <li>• Insufficiently monitoring/maintaining AI system</li> </ul>



# Potential harms and losses



# Emerging legal landscape

## Illustrative

### Intellectual property

- ***Kadrey et al. vs. Meta Platforms***  
Alleged use of copyrighted books in dataset used to train LLaMA

### Unfair competition

- ***Getty Images vs. Stability AI***  
Alleged unfair competitive advantage; alleged intellectual property infringement to build a competing business

### Data privacy

- ***P.M. et al. vs. Open AI***  
Alleged scraping of personally identifiable information without informed consent or knowledge

### Libel

- ***Walters vs. OpenAI***  
Alleged libel; spreading of misinformation

### Emerging IP legal questions with limited precedent:

1. Does a generative AI model trained on copyrighted material constitute infringement or fair use?
2. When are outputs generated with the assistance of generative AI systems considered human- vs. AI-generated?
3. When can generative AI prompts be protected under intellectual property law?

# Technological convergence — generative AI and beyond

Coming around the corner...

5G | 6G | edge computing

3-D reconstruction | mixed reality

Artificial intelligence | machine learning

Blockchain technology

Cloud and hosting

Semiconductors | graphic processing units



# Takeaways

When thinking about generative AI

1

Generative AI is technology built by humans, and prone to human error

2

Many risks are familiar, but new risks may arise as generative AI systems develop emergent capabilities or are applied to new use cases

3

AI laws and regulations are nascent and evolving, with implications for insurance



The information contained in this document, including but not limited to any recommendations, analysis or advice contained herein (the “Marsh Analysis”) provides only a general overview of subjects covered, is not intended to be taken as advice regarding any individual situation, and should not be relied upon as such. The information contained herein is based on sources we believe reliable, but we make no representation or warranty as to its accuracy. Marsh shall have no obligation to update the Marsh Analysis and shall have no liability to you or any other party arising out of this publication or any matter contained herein. Marsh makes no representation or warranty concerning the application of policy wording or the financial condition or solvency of insurers or reinsurers. Marsh makes no assurances regarding the availability, cost, or terms of insurance coverage. Insureds should consult their insurance and legal advisors regarding specific coverage issues. All insurance coverage is subject to the terms, conditions, and exclusions of the applicable individual policies. Marsh cannot provide any assurance that insurance can be obtained for any particular client or for any particular risk.

A business of Marsh McLennan