# TRAVELERS
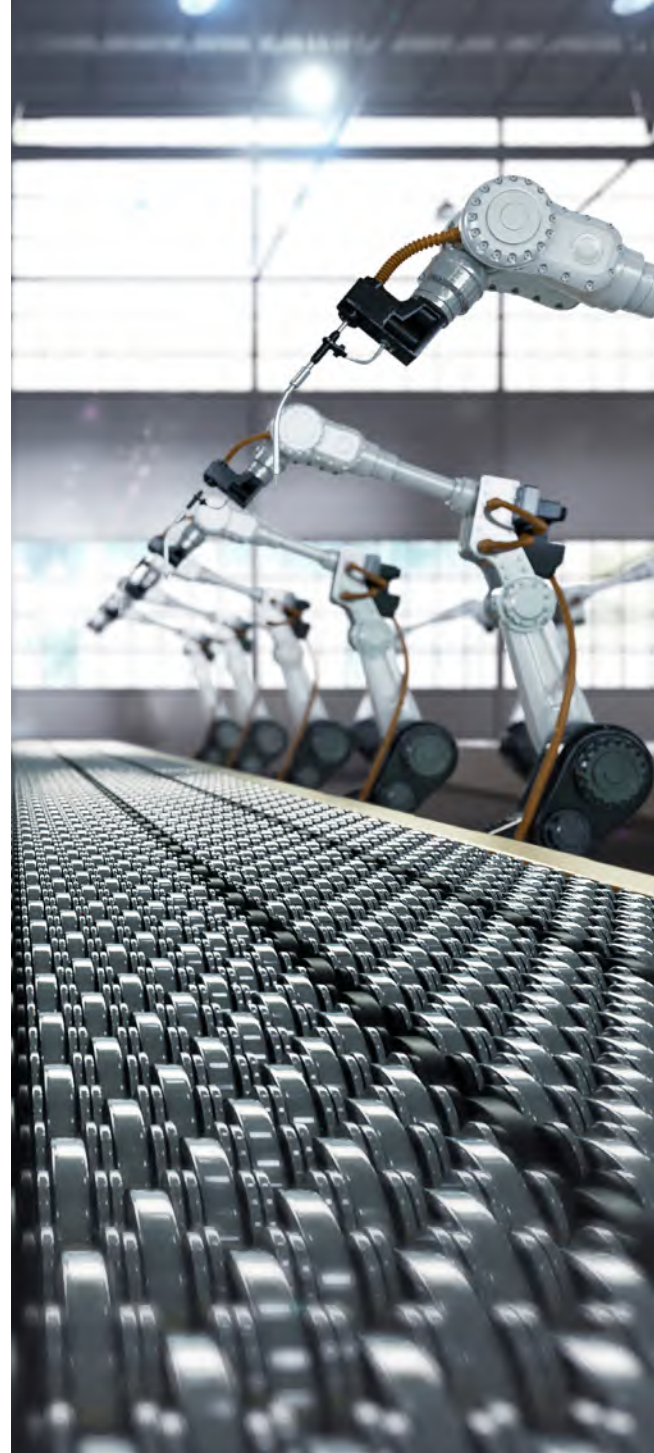
# Pursuing AI's Opportunities While Mitigating Its Risks

# Table of contents

# The State of AI: The Opportunity and the Risk

In artificial intelligence (AI), the world sees endless potential across various disciplines. For example, AI promises to change how doctors diagnose disease, law enforcement fights crime and autonomous vehicles navigate streets.

AI, once hailed as the key to technical innovations only seen in sci–fi novels, is quickly becoming a cost of entry into the competition for market advantage. Many companies know that they must invest in AI or risk extinction. While they aim to capitalize on AI's opportunity, many don't track the risks of using it.

"AI has captured our imaginations. It is expansively discussed but often misunderstood. Before leaders explore its potential, they must first understand how it works," says Mike Thoma, VP and Technology and Life Sciences Industry Lead at Travelers. "Armed with that knowledge, they should consider its implications – especially its risks. Only with their eyes wide open can they launch initiatives that best take advantage of AI."

## AI BY THE NUMBERS

*IN BUSINESS, RESEARCH AND LEGISLATION*

- **$790B+** by 2025 – The forecast of worldwide revenues for the AI market, including software, hardware and services.[1]

- **$15.7T** – AI's potential contribution to the global economy by 2030.[2]

- **5X** – Increase in research publications on fairness and transparency in AI between 2014 and 2021.[3]

- **486** – Mentions of AI in the Congressional Record of the 116th U.S. Congress (up from just 16 mentions in the 114th Congress).[4]

# AI: What It Is and How It Works



## WHAT IS AI?

With all the hype around AI, you may wonder: What, exactly, is it? When talking about AI, we're talking about a broad set of digital engineering techniques that help computer programs learn. These AI programs approximate the decision-making and problem-solving that come naturally to the human mind.

These programs are meant to improve processes, speed up tasks and enable better decisions. In short, AI aims to reduce human effort and improve outcomes. AI can accomplish repetitive, low-skilled tasks as well as complex work such as reading thousands of legal contracts in minutes while extracting information faster and with fewer errors than lawyers, paralegals and others involved in a law practice can.

## HOW DOES AI WORK?

AI performance has improved dramatically over the past decade as increasingly cheap and accessible compute and storage solutions have allowed vast quantities of data to be collected and leveraged. This data foundation has allowed advanced algorithms to create models suited to accomplishing specific tasks or business goals by repeatedly passing through a "training loop." Through this training process, the algorithm identifies correlations within the data set, especially any patterns that are usefully aligned with the desired task. For some model architectures, these relationships can be made interpretable to human insight, but in other cases it can be very difficult or even impossible to understand the algorithm's core functionality.

The algorithm selected will differ according to the underlying use case, the type and quantity of available data, and other considerations. Common examples include natural language processing (NLP) models, which are adept at handling free-form texts, deep learning image systems capable of classification and other imagery related tasks, and generalized linear models (GLMs), which have an inherent simplicity and explainability that allow them to be used in highly regulated use cases.

There is a lot of variability within AI solutions, but every case will have some broad commonalities. The first step is data ingestion, which may optionally also apply some normalization and cleaning processes to ensure that the data is appropriately prepared. Next is an inference step where the model's decision-making power is brought to bear on the ingested data. This is the most variable step depending on the type of AI algorithm used, but it is also where all the fascinating, and potentially inexplicable, patterns identified during the training stage are computationally leveraged to interpret the available data. The last step of an AI solution is implementation, where the mathematical output of the model is mapped to the business use case in such a way as to usefully complete some task or goal.

# Where AI Has Been and Where It's Going

The first technology that we might classify as artificial intelligence used rules created by humans. The early AI programs then used these rules to make decisions based on their programming.

Today's AI creates its own rules by being fed data from humans. The rules aim to re-create the same decisions humans would have made.

The next step for AI brings systems that can understand the context of their decisions and abstract concepts, adapt to new scenarios and explain their reasoning.

## TIMELINE: A BRIEF HISTORY OF AI[5]

*1950* – Alan Turing writes "Computing Machinery and Intelligence," suggesting machines can use information and logic to make decisions and solve problems, or, simply put, machines can think.

*1956* – The term "artificial intelligence" is coined at a conference hosted by Dartmouth College. During the conference, Logic Theorist, a program considered by many to be the first with artificial intelligence, is presented.

*1957–1973* – AI makes strides. Computer storage and processing speed improve, as do machine learning algorithms. Interest and research funding remain robust.

*1974–1980* – The first "AI winter," during which funding and research are slow, hindered by the processing speeds and computer memory available.

*1981–1986* – Japan pours funds into its Fifth Generation computer project, intending to become a world leader in computing technology. In response, the U.S. and Britain amp up their research efforts toward expert systems, which are used in bank loan screening and medical and sales applications.

*1987–1993* – Desktop computers overtake expert systems, and funding for AI research dips, heralding the second AI winter.

*THE 1990s* – Despite the second AI winter, statistical models rise for NLP, and the field focuses on intelligent agents – also called bots.

*1997* – IBM's Deep Blue beats chess grandmaster Garry Kasparov.

*2011* – IBM's Watson wins *Jeopardy!*, beating two reigning human champions.

*MODERN DAY* – AI goes mainstream as generative AI has taken the world by storm with the ability to create new content, including audio, video or text.

- Accurate NLP spurs market development of digital voice assistants in consumer smartphones, vehicles and homes.
- Emerging algorithm developments and growing data sets worldwide empower tech companies to use facial recognition to identify people with the same accuracy as humans.
- Along with these fast-moving AI advancements, privacy and ethical concerns gain steam.
- Notable progress made toward broader use of autonomous vehicles.

## PRESENT–DAY AI

Before recent AI advancements, the business world struggled to turn cumbersome yet valuable big data sets into improved productivity and innovation. But modern–day AI offers the capability to process and use a never–ending, ever–growing data supply.

Today, AI–based technologies can help businesses save time and money. Many already use it to automate processes, personalize marketing efforts and offer 24/7 customer service.

No matter who you are, you likely already interact with AI. It may even shape your day–to–day life in some way. We encounter it in:

- Digital assistants that answer spoken questions.

- Automatic speech–to–text transcription.

- Chatbots that provide customer service.

- Facial detection in photo software.

- Auditing software that detects credit card fraud.

- Newsfeeds that determine the headlines we see.

- Medical imaging analysis that helps healthcare professionals make decisions.

- Relevant product recommendations from your favorite online stores.

AI applications show the most promise when they solve a pressing need or vastly improve an experience either internally (for the business) or externally (for the customer).

Yet, while business leaders forge new paths in pursuit of these opportunities, they must remember that AI comes with new risks.

> *"Business and technology leaders must build a strong foundational strategy for AI that acknowledges its risks and plans for their mitigation. Their risk strategy helps set the groundwork to move AI forward with confidence."*
>
> Mike Thoma, VP and Technology and Life Sciences Industry Practice Lead, Travelers

# Two Sides of the Same Coin: AI's Opportunity Comes with Risks

## AI RISKS

AI can be a powerful tool, but it can also be unpredictable and cause unintended consequences if not used with caution. As with any powerful technology, there are potential negative impacts and unintended consequences that must be carefully considered. Risk identification is an essential early step of any successful AI project. The National Institute of Standards and Technology's (NIST) AI Risk Management Framework identifies these categories of risks associated with AI systems:

1. *ACCOUNTABILITY RISKS:* AI systems making decisions or recommendations that could be influenced by an organization's financial or business interests.

2. *SAFETY RISKS:* AI systems producing unintended results, leading to property damage, injury or death.

3. *RELIABILITY RISKS:* System malfunction or failure to meet performance requirements.

4. *BIAS RISKS:* Unintentional bias in decision-making due to lack of diversity in data or incorrect data labeling.

5. *SECURITY RISKS:* Unauthorized access, modification or destruction of AI systems and their data.

6. *PRIVACY RISKS:* Unauthorized access to private data or use of data for unintended purposes.

7. *EXPLAINABILITY RISKS:* Uncertainty in the decisions made by AI systems and lack of understanding of the AI system's decision process.

**AI risks can be different or increased compared to traditional software risks.**

According to the NIST AI Risk Management Framework, AI risks that are different or increased compared to traditional software include things like the potential for incorrect outcomes due to faulty data sets, the complexity of dealing with unanticipated outcomes, and the possibility of unexpected and potentially harmful interactions between different AI systems. It's important for technology companies to understand these risks and develop strategies to mitigate them.

**With all that goes into creating and training AI, there can be unintended consequences**

It's nearly impossible to foresee and plan for all outcomes. To move AI forward wisely, we must learn what we can from past mistakes. After all, AI has made headlines for notable failures.

### *AI'S NEAR-RUINOUS HOME PRICE PREDICTIONS*

Many digital companies that are focused on real estate use algorithms to predict home prices. Zillow took this further, using an AI program to inform Zillow Offers. Through this house-flipping program, Zillow purchased 27,000 homes starting in April 2018, but by September 2021, it had sold only 17,000 of those homes.[6] The algorithm's error rate was higher than Zillow realized, leading the company to lose $881 million in 2021.[7]

### *STRONG NEGATIVE REACTIONS TO FACIAL RECOGNITION SOFTWARE*

The Internal Revenue Service (IRS) faced widespread criticism when it used a commercial facial-recognition tool called ID.me to verify taxpayers' identities online. Some were concerned about privacy risks based on its access to users' biometric data. Other critics cited evidence that these systems can include deeply embedded gender and racial bias. The IRS halted its use of the program in early 2022 following a request sent by Senate Finance Chair Ron Wyden.[8]

As these examples demonstrate, AI's financial and reputational risks can come from various root causes, especially inadequately trained AI, when models are created without enough data points or scenarios.
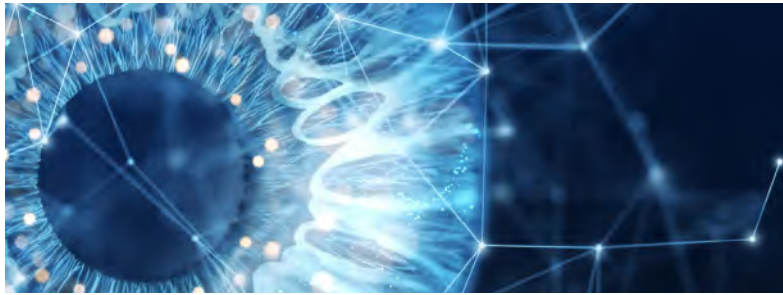
## THE COURT OF PUBLIC OPINION

Leaders who fail to mitigate risk while pursuing AI efforts may face public criticism and reputational harm — even costly investigations and lawsuits. Collectively, people still aren't sure what to think about AI. On the one hand, it's exciting. On the other, it may lead to new risks. According to Pew, the public doesn't understand AI and the majority only feel comfortable with it when humans retain control of final decisions and actions.

Many support defining higher standards for testing and developing high-risk AI applications such as driverless cars and brain chip implants.[9]

"In general, people hold machines to a higher level than humans," noted Jen Klein, Underwriting Officer and Information Technology Lead at Travelers. "With human operators the public typically tolerates, even expects, errors within a certain range. But machine-made errors are less acceptable. For example, if a programming error in an autonomous vehicle platform causes an accident, the technology is held to a higher standard than a human operator would be if they made the same error and caused an accident. Navigating the court of public opinion is essential when it comes to successfully implementing AI."

# Reducing Risks Associated with AI

It is important for technology companies to exercise caution when developing and using AI to ensure the safety of their products and services. By taking these proactive steps described within NIST's AI Risk Management Framework, technology companies can help mitigate the risks of AI:

- **Proper design and programming:** Ensure that the technology is designed and programmed properly. This includes ensuring that the AI system has clearly defined goals and constraints, and that it is programmed to follow any applicable regulations.

- **Data quality:** Confirming data quality is a critical part of reducing bias in AI systems. High-quality data is essential to getting accurate results. Additionally, it is important to thoroughly analyze the data to ensure that it meets the required standards and that it can be used to create the most accurate results.

- **Human oversight:** Another way to mitigate the risks of AI is to incorporate human oversight into the design and use of the technology. This could involve having AI systems make decisions that are subject to human approval.

- **Transparency and accountability:** Making AI systems transparent and accountable can also help to mitigate their risks. This could involve making the algorithms and data used by AI systems available for review and audit, as well as establishing clear guidelines for the use of the technology.

- **Continual testing and evaluation:** Continually testing and evaluating AI systems helps ensure that they are functioning properly and not causing any harm. This could involve conducting regular audits and assessments, as well as implementing feedback mechanisms to identify and address any issues that arise.

## UNDERSTAND THE IMPORTANCE OF AI EXPLAINABILITY[10]

When people outside of the programming team want to know about your AI, they'll usually be interested in how it makes its decisions. This is an idea called "explainability." In some situations it may be important to understand and be able to explain your AI. Otherwise, you may face unknown risks. There are cases where even the engineers designing AI systems don't understand how their programs work. Don't allow this to be the case in your organization.

The most essential aspect of explainability is to develop a culture of governance. Meticulously documenting every step in the design and training of an AI improves your ability to explain how and why it gets the results it does.

Explainability can include details on the design, operation and limitations of the AI program. Explainability may also include:

- What the system is for.

- What it is not for.

- How it was designed.

- What its limitations are.

Legislation and various regulatory agencies may have requirements for AI levels of explainability. These requirements can include specifics about the privacy, security and algorithmic transparency of AI. You may be required to share where the AI's data comes from and how it's used, which can also improve public acceptance of AI.

## PAY ATTENTION TO WHERE AI DECISIONS GET MADE

When AI makes its own decisions, leaders must dig deeper. The further away the AI gets from a human making the decision, the more serious its risks. It's essential to responsibly develop and define how any AI program is supervised by humans.

The suggested level of human moderation depends on several factors, including the purpose and potential consequences of the system. For example, a higher level of supervision is required for higher-risk AI.

*"There are generally lower levels of risk when a human expert is in control of decisions. In the middle, humans are monitoring AI systems and can intervene if needed. Higher levels of risk exist when AI makes decisions on its own."*

Jen Klein, Underwriting Officer and Information Technology Underwriting Lead, Travelers

# WHO'S RESPONSIBLE FOR AI RISKS?

Many companies use AI solutions provided by a third-party vendor to reduce their AI risks. They may assume that because they're not creating the AI in-house, they aren't responsible for potential mistakes or bad outcomes. On the contrary — in many cases, businesses that buy AI-enabled tools are still accountable for the programs, their outcomes and their effects. Additionally, AI vendors may want to share very little about their programs' inner workings. If this is your situation, you may be more exposed to liability than you realize.

Consider these potential AI use cases in electronics manufacturing and life sciences — and the issues that crop up while employing them.
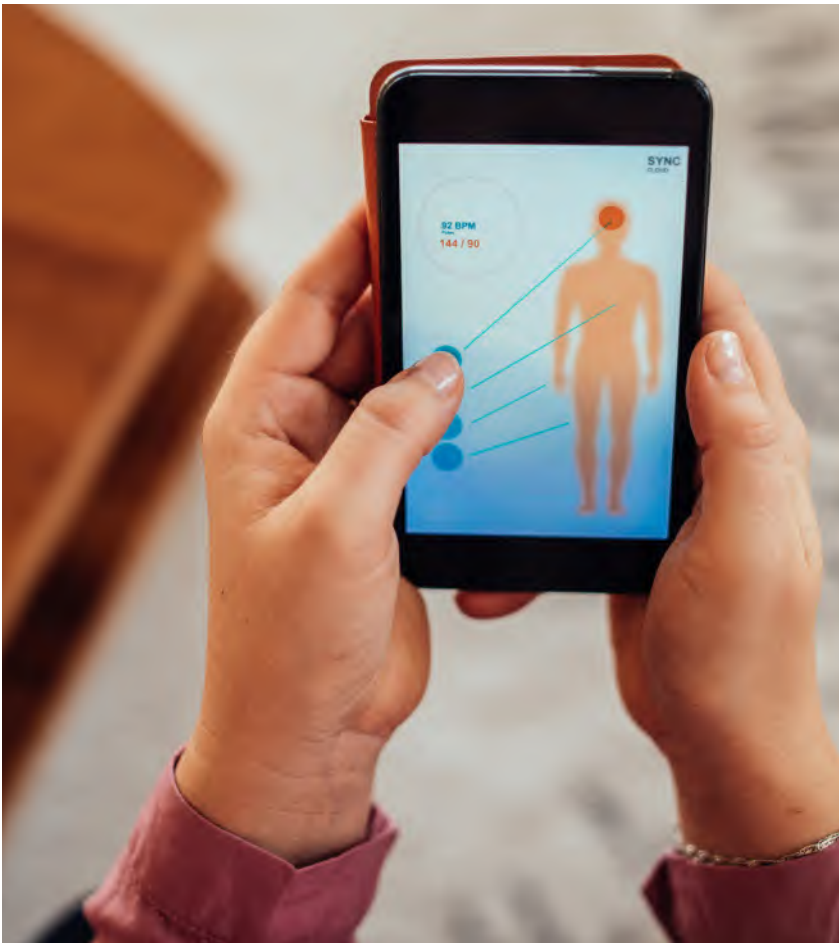


## SCENARIO 1 – MEDICAL DIAGNOSIS

Medical professionals and AI developers team up to use AI for diagnosing health conditions. The technology processes imaging test results, like X-rays, to screen for anything from dental cavities to cancer. The AI may diagnose some diseases earlier, faster and more precisely than human doctors. Yet, if improperly trained or tampered with, the AI could over- or underdiagnose patients, potentially in a biased way. Some diseases could be missed, while other patients may undergo unnecessary treatments.

## SCENARIO 2 — HEALTH AI IN WEARABLE TECHNOLOGY

Medical researchers develop AI using wearable technology that can detect cardiovascular symptoms. Yet if developers don't thoroughly validate and test their model, it could scare users with false alerts. Patients flood physicians' offices only to find that the alerts are erroneous, and complaints flood the wearable technology company's customer service department. Patients may eventually form a class action lawsuit seeking compensation for their time, stress and unnecessary medical testing.

# Implement Insurance and Contractual Risk Transfer Solutions

Beyond making every effort to build and employ AI ethically and responsibly, there are two main tools companies can use to mitigate remaining AI risk: contractual risk transfer and strategic insurance solutions.

## CONTRACTUAL RISK TRANSFER

Contractual risk transfer (CRT) is an important step that AI developers often overlook. It helps ensure that the liability risk is placed with the party most able to control it. A contractual risk management program can help organizations that develop AI better protect themselves, typically by reducing the financial exposures they may face. It does so by transferring liability related to their services and their partners' use of their products away from them and onto their partners or others.

Through CRT, AI developers clarify what they provide and who is responsible for what. They can use customer contracts to answer key questions in advance, including whether the AI is customized or off the shelf, how the AI makes decisions and what its end use will be.

It's vital to ensure that these contracts are well written and include proper language addressing liability. All technology contracts should include these three contractual provisions:

1. **LIMITATION OF LIABILITY.** This provision disclaims liability for certain types of damages and can provide a dollar threshold above which the AI developer will not be contractually responsible. These typically include incidental, consequential and special damages such as lost profits, sales and data.

2. **LIMITATION OF WARRANTIES.** Such a provision identifies each type of warranty and what it does and does not cover. It also identifies the remedies provided by the warranties if the product or work does not comply.
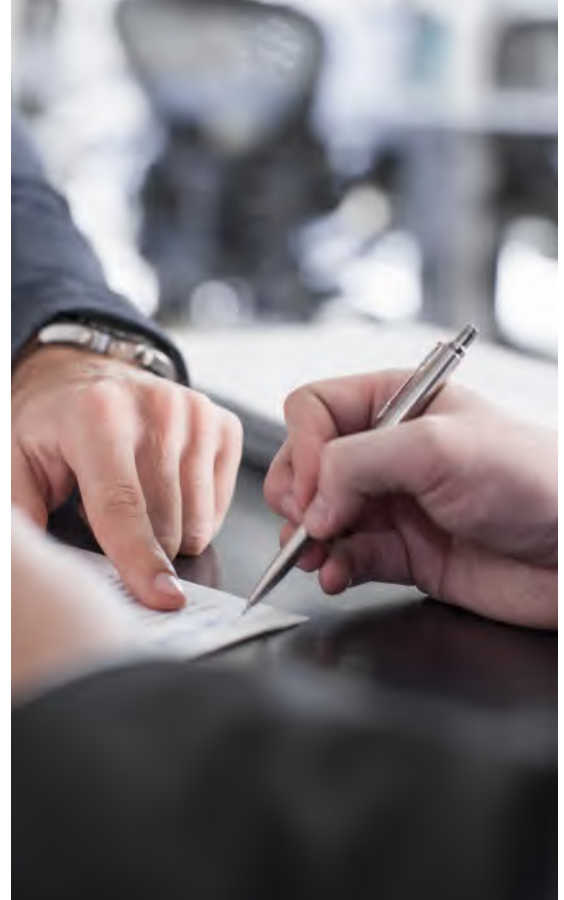
3. **LIMITATION OF REMEDIES.** This provision specifically lays out the sole remedy available to a non-breaching party — typically a refund of the purchase price or replacement of the allegedly defective product or service.

Because AI developers often use CRT to shift responsibility for what they produce to the user, companies that outsource AI should pay close attention to their user agreements. These may contain disclaimers, damage limitations and liability waivers.

After procuring AI-enabled software and signing these legal documents, organizations should be aware of the new risks they've taken on so they can plan for any issues that could arise from using the product. Because software companies often limit damages to the original cost of the software, this may leave purchasers of AI-enabled products on the hook for most of the costs should anything go awry. That's where proper insurance coverage kicks in.

## INSURANCE SOLUTIONS FOR AI RISKS

The right insurance coverage can help safeguard technology companies that develop or use AI from third-party claims for damages or first-party loss. The primary types of liability coverages include technology errors & omissions, privacy and security (cyber), and commercial general liability. Privacy breach notification expense, a first-party coverage, can also come into play.

## LIABILITY INSURING AGREEMENTS:

### Technology Errors & Omissions

This insurance provides coverage for claims that arise from errors, omissions or negligent acts by the insured in the performance of or failure to perform work or services, or the failure of products to function as intended. It can help protect your business from the costs of covered claims that might arise from AI model glitches, improper AI training or poor-performing AI that fails to work as intended.

### Privacy and Security

This insurance provides coverage for claims for damages that arise out of data or privacy breaches as well as other alleged cyber events. It can help protect your business from the cost of claims that may arise from the following:

- Failure to prevent the unauthorized access to private or confidential information.

- Failure to comply with a privacy policy.

- Wrongful collection of private or confidential information.

- Failure to provide notification of a data breach where required by law.

- Failure to destroy confidential information.

- Failure to prevent a security breach.

### Commercial General Liability Insurance

When AI-driven software in a business process accidentally causes bodily injury or property damage, Commercial General Liability (CGL) may come into play. CGL helps protect your business from the costs of claims when your products or your work cause bodily injury or property damage to a third party.

## BREACH RESPONSE INSURING AGREEMENTS:

### Privacy Breach Notification

When organizations suffer a data breach, they must follow all applicable laws and regulations to notify affected individuals and sometimes provide services to help protect them from further harm. Costs to provide call center support, credit monitoring services, identity fraud insurance and notification all add up. This type of insurance helps your business cover those costs.

## WRAP-UP

AI has become more popular today than ever, yet some may be surprised to know the ideas for AI were first developed over 65 years ago in 1956. Tech companies that have a better understanding of the risks associated with their AI systems can make more informed decisions on how best to proceed with its development and deployment.

Investment in creating effective risk management systems to identify, assess and monitor AI systems can help prepare for future AI-driven decision-making. Managing the risks of AI along the way is just as important as exploring possibilities and opportunities.

# AI: A Glossary of Terms

## A GLOSSARY OF AI TERMS

*MODEL* – Once a machine-learning algorithm has been trained on data, the output of the process is known as the model. This can then be used to make predictions.

*ALGORITHM* – Generally a set of rules and steps used to solve a problem or complete a task. When used in AI, algorithms operate within a model that guides computer programs to learn and operate autonomously.

*DATA SET* – A collection of data points used to train AI algorithms.

*TRAINING* – The process of giving AI the information (data set) it needs to learn, perform tasks and produce outcomes. Depending on the type of AI, training involves collecting data to analyze using algorithms. During training, AI learns how to interpret and use this data properly. Before completing AI training, developers should complete a validation phase with a new data set and a testing phase to ensure proper function and accuracy.

*EXPLAINABILITY* – The ability of AI data scientists to explain how an AI model performs tasks, comes to its conclusions or makes decisions. Sometimes this process is opaque to the data scientists.

*MACHINE LEARNING (ML)* – This is a subset of AI that often uses statistical models to enable computer systems with the ability to "learn" from data without being explicitly given the instructions for how to do so.

*DEEP LEARNING (DL)* – This is a subset of machine learning that attempts to mimic the neural networks found in the brain. Using a large number of layers helps enable the model to learn how to recognize complex patterns in data.

*NATURAL LANGUAGE PROCESSING (NLP)* – NLP allows digital technology to understand and respond to human language through spoken word or text input.

# About Travelers

The Travelers Companies, Inc. (NYSE: TRV) is a leading provider of property casualty insurance for auto, home and business. A component of the Dow Jones Industrial Average, Travelers has approximately 30,000 employees and generated revenues of approximately $37 billion in 2022.

*For more information, visit travelers.com.*

**TRAVELERS**

# SOURCES

[1] International Data Corporation (IDC). *Worldwide Artificial Intelligence Software Forecast, 2022-2026*. Feb. 15, 2022.
https://www.idc.com/getdoc.jsp?containerId=US49571222

[2] PwC. *Global Artificial Intelligence Study: Exploiting the AI Revolution: What's the real value of AI for your business and how can you capitalize?*
https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html

[3] Stanford University Human-Centered Artificial Intelligence (HAI), not Computing Research Association & AI Index. *Artificial Intelligence Index Report 2021*. https://hai.stanford.edu/ai-index-2021

[4] Stanford University Human-Centered Artificial Intelligence (HAI), not Computing Research Association & AI Index. *Artificial Intelligence Index Report 2021*. https://hai.stanford.edu/ai-index-2021

[5] Dataversity. *A Brief History of Artificial Intelligence*. Jan. 17, 2022.
https://www.dataversity.net/brief-history-artificial-intelligence/.

[6] CNN Business. *Zillow's home-buying debacle shows how hard it is to use AI to value real estate*. Nov. 9, 2021.
https://www.cnn.com/2021/11/09/tech/zillow-ibuying-home-zestimate/index.html

[7] Wall Street Journal. *Zillow's Shuttered Home-Flipping Business Lost $881 Million in 2021*. Feb. 10, 2022.
https://www.wsj.com/articles/zillows-shuttered-home-flipping-business-lost-881-million-in-2021-11644529656

[8] United States Senate Committee on Finance. *Wyden Calls on IRS To End Use of Facial Recognition for Online Accounts*. Feb. 7, 2022.
https://www.finance.senate.gov/chairmans-news/wyden-calls-on-irs-to-end-use-of-facial-recognition-for-online-accounts

[9] Pew Research Center. *AI and Human Enhancement: Americans' Openness Is Tempered by a Range of Concerns*. March 17, 2022.
https://www.pewresearch.org/internet/2022/03/17/ai-and-human-enhancement-americans-openness-is-tempered-by-a-range-of-concerns/

[10] Government Accountability Office. *Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities*. June 2021.
https://www.gao.gov/assets/gao-21-519sp.pdf