INFO~TECH RESEARCH GROUP

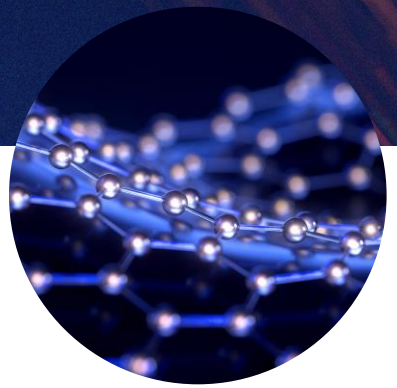
Michael Best

AI Trends Report 2023

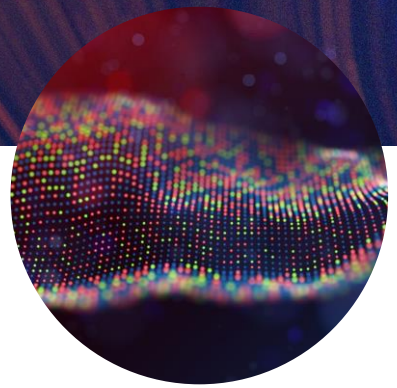
INFO~TECH RESEARCH GROUP

Michael Best

Design for AI

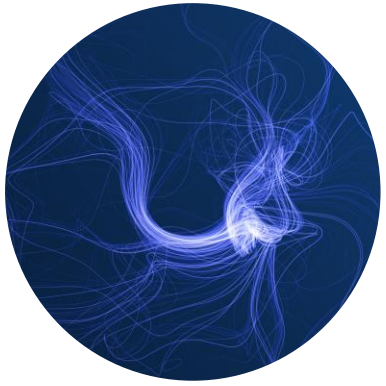


Event-Based Insights



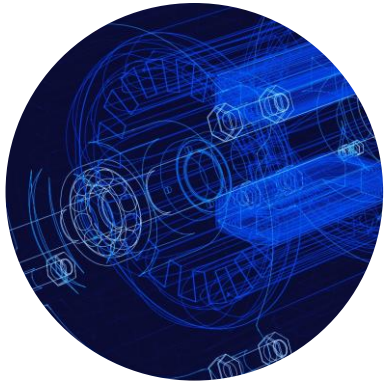
Synthetic Data

Edge AI

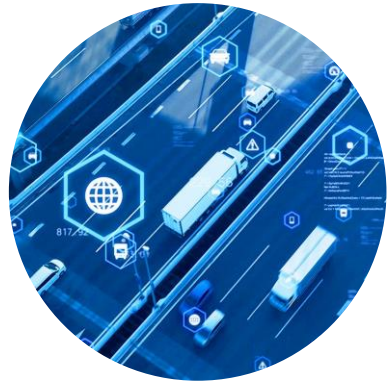


AI in Science and Engineering

AI Reasoning



Digital Twin



Combinatorial Optimization

Challenges that slowed the adoption of AI

Data Readiness

- Lack of unified systems and unified data
- Data quality issues
- Lack of the right data required for machine learning

ML Operations Capabilities

- Lack of tools, technologies, and methodologies to operationalize models created by data scientists

Understanding of AI Role and Its Business Value

- Lack of understanding of AI use cases – how AI/ML can be applied to solve specific business problems
- Lack of understanding how to define the business value of AI investments

To overcome the challenges, enterprises adopted different strategies

- Improve data management capabilities, including data governance and data initiatives
- Create data catalogs
- Document data and information architecture
- Solve data-related problems including data quality, privacy, and ethics

- Increase availability of cloud platforms, tools, and capabilities
- Develop and grow machine learning operations (MLOps) tools, platforms, and methodologies to enable model operationalizing and monitoring in production

- Identify AI C-suite toolkits (for example, Empowering AI Leadership from the World Economic Forum, 2022)
- Document industry use cases
- Use frameworks and tools to define business value for AI investments

Design for AI

Sustainable AI system design needs to consider several aspects: the business application of the system, data, software and hardware, governance, privacy, and security.

It is important to define from the beginning how AI will be used by and for the application to clearly articulate business value, manage expectations, and set goals for the implementation.

Design for AI will change how we store and manage data and how we approach the use of data for development and operation of AI systems.

An AI system design approach should cover all stages of AI lifecycle, from design to maintenance. It should also support and enable iterative development of an AI system.

To take advantage of different tools and technologies for AI system development, deployment, and monitoring, the design of an AI system should consider software and hardware needs and design for seamless and efficient integrations of all components of the system and with other existing systems within the enterprise.

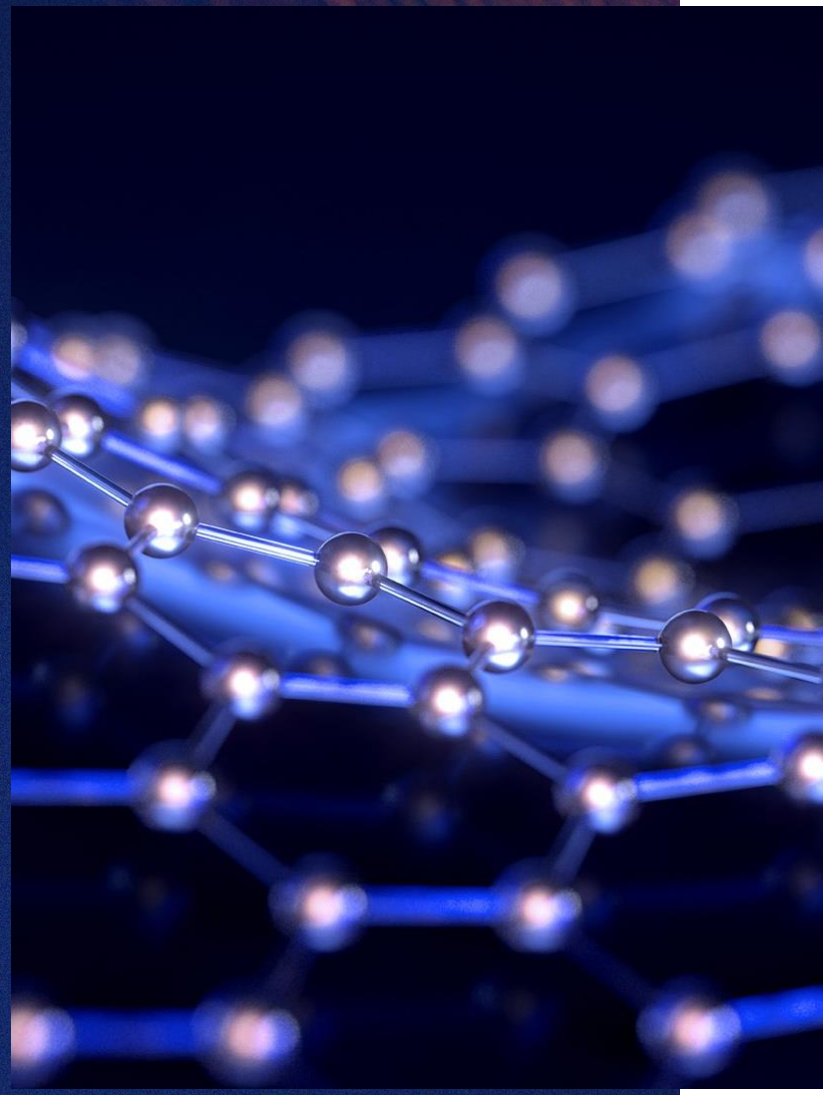
AI in Science and Engineering

AI helps sequence genomes to identify variants in a person's DNA that indicate genetic disorders. It allows researchers to model and calculate complicated physics processes, to forecast the genesis of the universe's structure, and to understand planet ecosystem to help advance the climate research. AI drives advances in drug discovery and can assist with molecule synthesis and molecular property identification.

AI finds application in all areas of science and engineering. The role of AI in science will grow and allow scientists to innovate faster.

AI will further contribute to scientific understanding by assisting scientists in deriving new insights, generating new ideas and connections, generalizing scientific concepts, and transferring them between areas of scientific research.

Using synthetic data and combining physical and machine learning models and other advances of AI/ML – such as graphs, use of unstructured data (language models), and computer vision – will accelerate the use of AI in science and engineering.



Event- and Scenario-Driven AI

AI-driven signal-gathering systems analyze a continuous stream of data to generate insights and predictions that enable strategic decision modeling and scenario planning by providing understanding of how and what areas of business might be impacted by certain events.

AI enables the scenario-based approach to drive insights through pattern identification in addition to familiar pattern recognition, helping to understand how events are related.

A system with anticipatory capabilities requires an event-driven architecture that enables gathering and analyzing different types of data (text, video, images) across multiple channels (social media, transactional systems, news feeds, etc.) for event-driven and event-sequencing modeling.

ML simulation-based training of the model using advanced techniques under the umbrella of Reinforcement Learning in conjunction with statistically robust Bayesian probabilistic framework will aid in setting up future trends in AI.

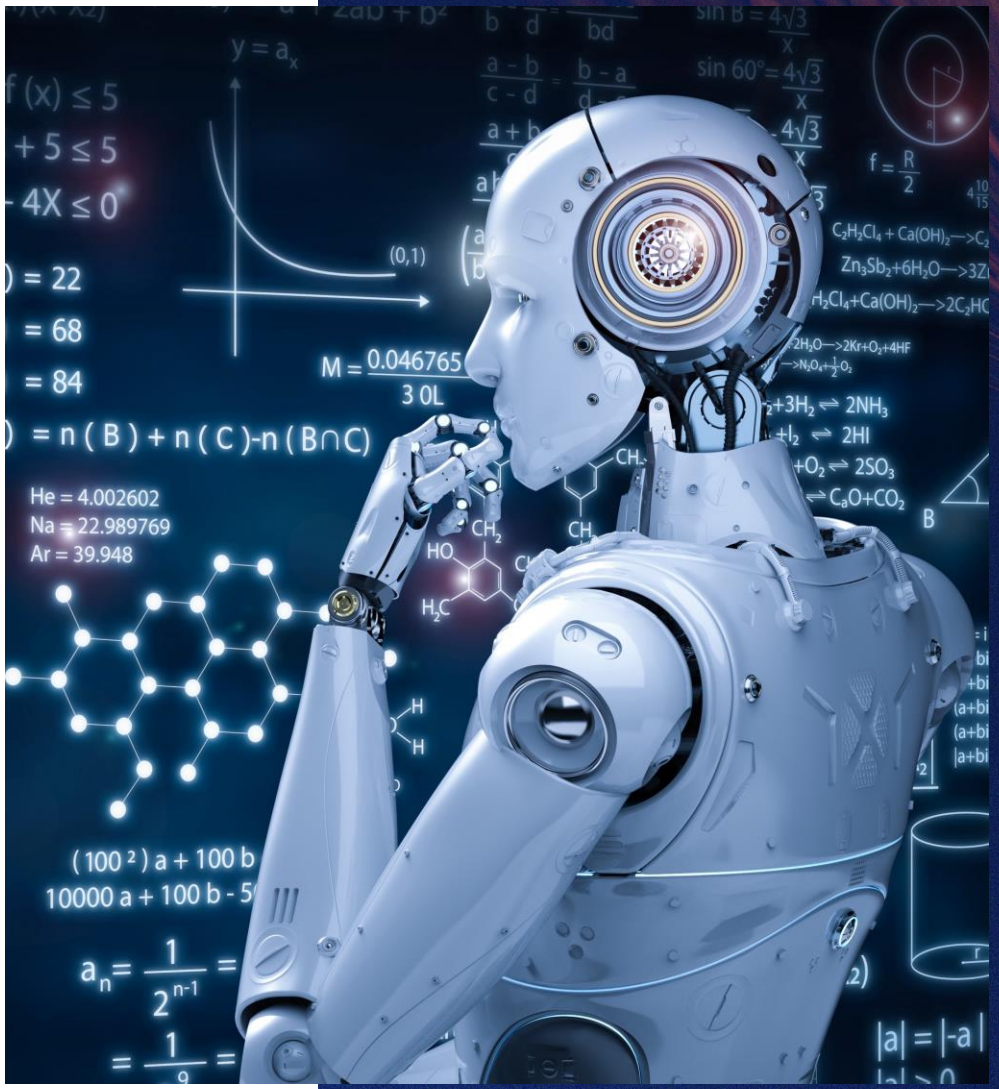
AI Reasoning

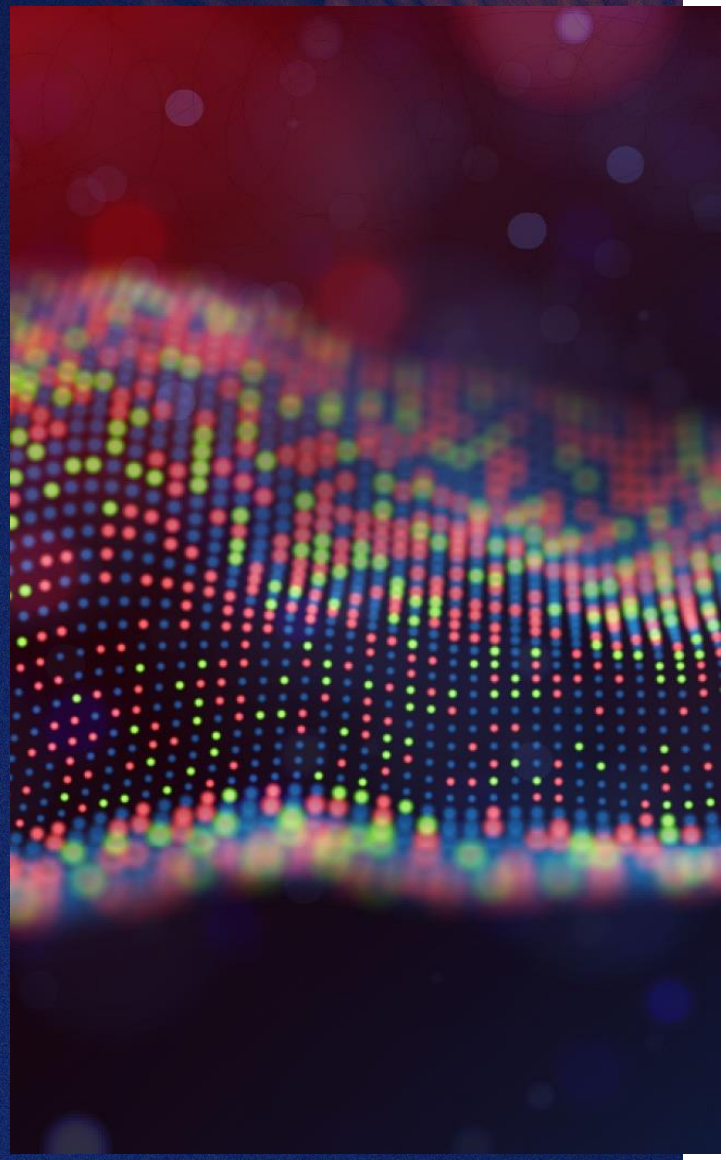
Most of the applications of machine learning and AI today is about predicting future behaviors based on historical data and past behaviors. We can predict what product the customer would most likely buy or the price of a house when it goes on sale.

Most of the current algorithms use the correlation between different parameters to make a prediction, for example, the correlation between the event and the outcome can look like "When X occurs, we can predict that Y will occur." This, however, does not translate into "Y occurred because of X."

The development of a causal AI that uses causal inference to reason and identify the root cause and the causal relationships between variables without mistaking correlation and causation is still in its early stages but rapidly evolving.

Some of the algorithms that the researchers are working with are casual graph models and algorithms that are at the intersection of causal inference with decision making and reinforcement learning (Causal Artificial Intelligence Lab, 2022).





Synthetic Data

Synthetic data is artificially generated data that mimics the structure of real-life data. It should also have the same mathematical and statistical properties as the real-world data that it is created to replicate.

Synthetic data is used to train machine learning models when there is not enough real data or the existing data does not meet specific needs. It allows users to remove contextual bias from data sets containing personal data, prevent privacy concerns, and ensure compliance with privacy laws and regulations.

Another application of synthetic data is solving data-sharing challenges.

Researchers learned that quite often synthetic data sets outperform real-world data. Recently, a team of researchers at MIT built a synthetic data set of 150,000 video clips capturing human actions and used that data set to train the model. The researchers found that "the synthetically trained models performed even better than models trained on real data for videos that have fewer background objects" (MIT News Office, 2022).

Today, synthetic data is used in language systems, in training self-driving cars, in improving fraud detection, and in clinical research, just to name a few examples.

Synthetic data opens the doors for innovation across all industries and applications of AI by enabling access to data for any scenario and technology and business needs.

Digital Twins

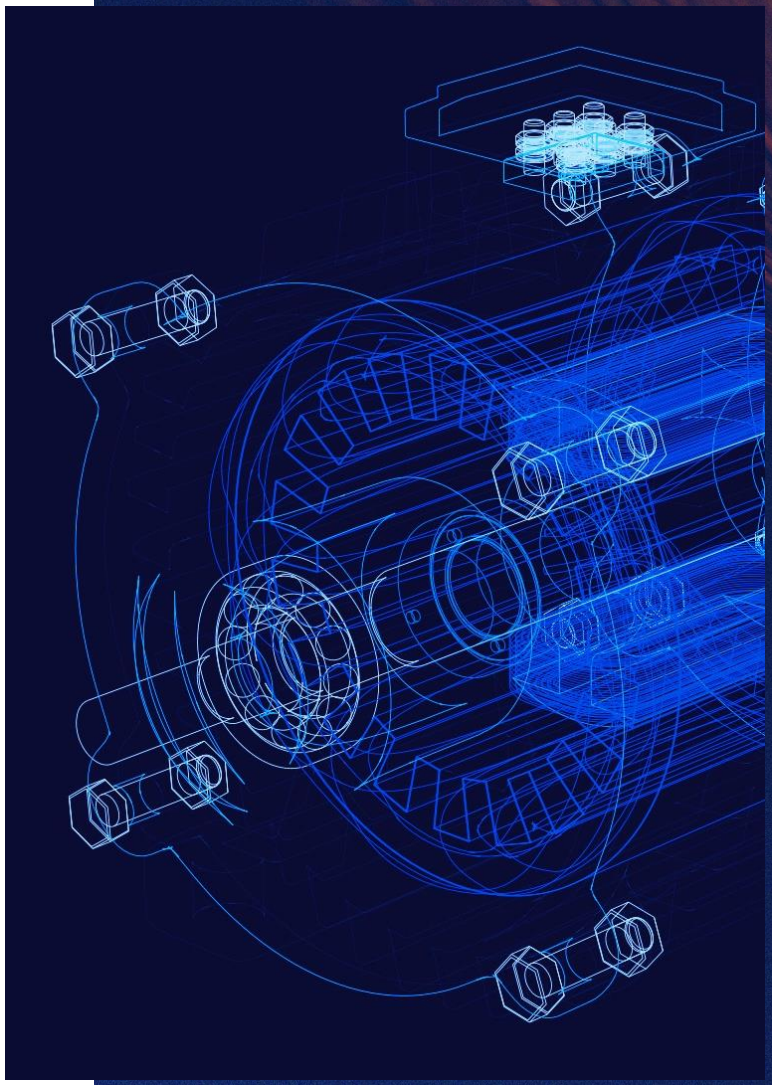
Digital twins (DT) are virtual replicas of physical objects, devices, people, places, processes, and systems. In Manufacturing, almost every product and manufacturing process can have a complete digital replica of itself thanks to IoT, streaming data, and cheap cloud storage.

All this data has allowed for complex simulations of, for example, how a piece of equipment will perform over time to predict future failures before they happen, reducing costly maintenance and extending equipment lifetime.

In addition to predictive maintenance, DT and AI technologies have enabled organizations to design and digitally test complex equipment such as aircraft engines, trains, offshore oil platforms, and wind turbines before physically manufacturing them. This helps to improve product and process quality, manufacturing efficiency, and costs. DT technology also finds applications in architecture, construction, energy, infrastructure industries, and even retail.

Digital twins combined with the metaverse provide a collaborative and interactive environment with immersive experience and real-time physics capabilities (as an example, Siemens presented an Immersive Digital Twin of a Plant at the Collision 2022 conference).

Future trends include enabling autonomous behavior of a DT. An advanced DT can replicate itself as it moves into several devices, hence requiring the autonomous property. Such autonomous behavior of the DT will in turn influence the growth and further advancement of AI.



Edge AI

A simple definition for edge AI: A combination of edge computing and artificial intelligence, it enables the deployment of AI applications in devices of the physical world, in the field, where the data is located, such as IoT devices, devices on the manufacturing floor, healthcare devices, or a self-driving car.

Edge AI integrates AI into edge computing devices for quicker and improved data processing and smart automation.

The main benefits of edge AI include:

- Real-time data processing capabilities to reduce latency and enable near real-time analytics and insights.
- Reduced cost and bandwidth requirements as there is no need to transfer data to the cloud for computing.
- Increased data security as the data is processed locally, on the device, reducing the risk of loss of sensitive data.
- Improved automation by training machines to perform automated tasks.

Edge AI is already used in a variety of applications and use cases including computer vision, geospatial intelligence, object detection, drones, and health monitoring devices.

Combinatorial Optimization

"Combinatorial optimization is a subfield of mathematical optimization that consists of finding an optimal object from a finite set of objects" (Wikipedia, retrieved December 2022).

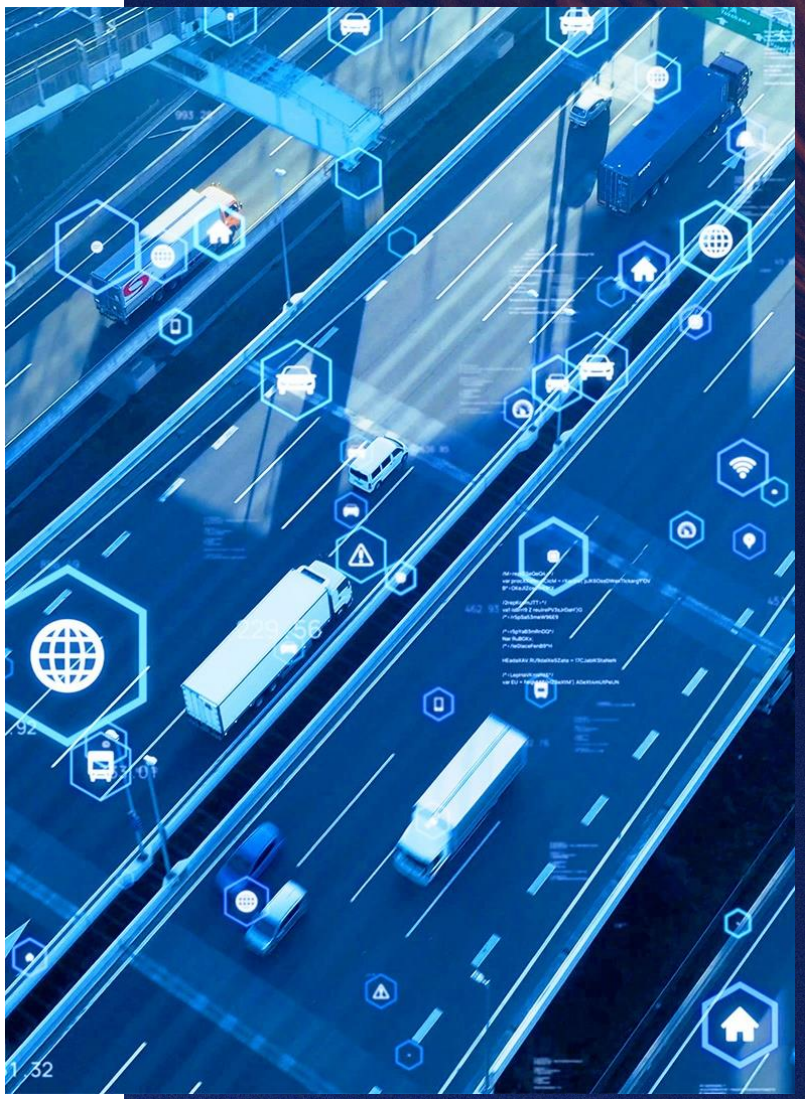
Applications of combinatorial optimization include:

- Supply chain optimization
- Scheduling and logistics, for example, vehicle routing where the trucks are making stops for pickup and deliveries
- Operations optimization

Classical combinatorial optimization (CO) techniques were widely used in operations research and played a major role in earlier developments of AI.

The introduction of deep learning algorithms in recent years allowed researchers to combine neural network and conventional optimization algorithms; for example, incorporating neural combinatorial optimization algorithms in the conventional optimization framework. Researchers confirmed that certain combinations of these frameworks and algorithms can provide significant performance improvements.

The research in this space continues and we look forward to learning how machine learning and AI (backtracking algorithms, reinforcement learning, deep learning, graph attention networks, and others) will be used for solving challenging combinatorial and decision-making problems.



References

"AI Can Power Scenario Planning for Real-Time Strategic Insights." *The Wall Street Journal, CFO Journal,* content by Deloitte, 7 June 2021. Accessed 11 Dec. 2022.

Ali Fdal, Omar. "Synthetic Data: 4 Use Cases in Modern Enterprises." *DATAVERSITY*, 5 May 2022. Accessed
11 Dec. 2022.

Andrews, Gerard. "What Is Synthetic Data?" *NVIDIA*, 8 June 2021. Accessed 11 Dec. 2022.

Bareinboim, Elias. "Causal Reinforcement Learning." *Causal AI*, 2020. Accessed 11 Dec. 2022.

Bengio, Yoshua, Andrea Lodi, and Antoine Prouvost. "Machine learning for combinatorial optimization: A methodological tour d'horizon." *European Journal of Operational Research*, vol. 290, no. 2, 2021, pp. 405-421, https://doi.org/10.1016/j.ejor.2020.07.063. Accessed 11 Dec. 2022.

Benjamins, Richard. "Four design principles for developing sustainable AI applications." *Telefónica S.A.,* 10 Sept. 2018. Accessed on 11 Dec. 2022.

Blades, Robin. "AI Generates Hypotheses Human Scientists Have Not Thought Of." *Scientific American,* 28 October 2021. Accessed 11 Dec. 2022.

"Combinatorial Optimization." Wikipedia article, Accessed 11 Dec. 2022.

Cronholm, Stefan, and Hannes Göbel. "Design Principles for Human-Centred Artificial Intelligence." University of Borås, Sweden, 11 Aug. 2022. Accessed on 11 Dec. 2022

Devaux, Elise. "Types of synthetic data and 4 real-life examples." Statice, 29 May 2022. Accessed 11 Dec. 2022.

Emmental, Russell. "A Guide to Causal AI." ITBriefcase, 30 March 2022. Accessed 11 Dec. 2022.

"Empowering AI Leadership: AI C-Suite Toolkit." World Economic Forum, 12 Jan. 2022. Accessed 11 Dec 2022.

Falk, Dan. "How Artificial Intelligence Is Changing Science." *Quanta Magazine*, 11 March 2019. Accessed 11 Dec. 2022.

Fritschle, Matthew J. "The Principles of Designing AI for Humans." Aumcore, 17 Aug. 2018. Accessed 8 Dec. 2022.

Garmendia, Andoni I., et al. Neural Combinatorial Optimization: a New Player in the Field." IEEE, arXiv:2205.01356v1, 3 May 2022. Accessed 11 Dec. 2022.

Gülen, Kerem. "AI Is Revolutionizing Every Field and Science is no Exception." Dataconomy Media GmbH, 9 Nov. 9, 2022. Accessed 11 Dec. 2022

Krenn, Mario, et al. "On scientific understanding with artificial intelligence." *Nature Reviews Physics*, vol. 4, 11 Oct. 2022, pp. 761–769. https://doi.org/10.1038/s42254-022-00518-3. Accessed 11 Dec. 2022.

Laboratory for Information and Decision Systems. "The real promise of synthetic data." *MIT News,* 16 Oct. 2020. Accessed 11 Dec. 2022.

Lecca, Paola. "Machine Learning for Causal Inference in Biological Networks: Perspectives of This Challenge." *Frontiers*, 22 Sept. 2021. Accessed 11 Dec. 2022.

References

Mirabella, Lucia. "Digital Twin x Metaverse: real and virtual made easy." Siemens presentation at Collision 2022 conference, Toronto, Ontario. Accessed 11 Dec. 2022.

Mitchum, Rob, and Louise Lerner. "How AI could change science." *University of Chicago News*, 1 Oct. 2019. Accessed 11 Dec. 2022.

Okeke, Franklin. "The benefits of edge AI." *TechRepublic*, 22 Sept. 2022, Accessed 11 Dec. 2022.

Perlmutter, Nathan. "Machine Learning and Combinatorial Optimization Problems." *Crater Labs*, 31 July 31, 2019. Accessed 11 Dec. 2022.

Sampson, Ovetta. "Design Principles for a New AI World." *UX Magazine*, 6 Jan. 2022. Accessed 11 Dec. 2022.

Sgaier, Sema K., Vincent Huang, and Grace Charles. "The Case for Causal AI." *Stanford Social Innovation Review*, Summer 2020. Accessed 11 Dec. 2022.

"Synthetic Data." Wikipedia article, Accessed 11 Dec. 2022.

Take, Marius, et al. "Software Design Patterns for AI-Systems." *EMISA Workshop 2021*, CEUR-WS.org, Proceedings 30. Accessed 11 Dec. 2022.

Toews, Rob. "Synthetic Data Is About To Transform Artificial Intelligence." *Forbes*, 12 June 2022. Accessed
11 Dec. 2022.

Zewe, Adam. "In machine learning, synthetic data can offer real performance improvements." *MIT News Office*, 3 Nov. 2022. Accessed 11 Dec. 2022.

Zhang, Junzhe, and Elias Bareinboim. "Can Humans Be out of the Loop?" Technical Report, *Department of Computer Science*, Columbia University, NY, June 2022. Accessed 11 Dec. 2022.

Contributors

Irina Sedenko
Advisory Director
Info-Tech

Anu Ganesh
Technical Counselor
Info-Tech

Jason Schultz
Chief Innovation and
Technology Officer
Michael Best

Sarah Alt
Chief Process
& AI Officer
Michael Best

Amir Feizpour
Co-Founder & CEO
Aggregate Intellect Inc.

David Glazer
VP of Analytics
Kroll

Delina Ivanova
Associate Director, Data & Analytics
HelloFresh

Usman Lakhani
DevOps
WeCloudData


INFO~TECH
RESEARCH GROUP
Michael
Best