

# Cutting through the noise: demystifying the buzz around artificial intelligence in financial decision-making

September 2020

When artificial intelligence (AI) first emerged in 1956, there was no way to predict the magnitude of its impact. AI is an underlying element of countless technologies with which we regularly interact, and that are now a part of everyday life – from email to ride sharing to banking to media streaming.

A recurring theme in the AI discourse is its potential to enhance or even replace human intelligence. At RepRisk, we believe that AI is not a replacement for human intelligence, but rather a way to further it. Specifically, supervised machine learning is the most effective mechanism to further human intelligence, as it produces more relevant and timely data – criteria that are critical in fields like the financial services sector.

The financial services sector was an early adopter of artificial intelligence, as they quickly recognized its potential to generate alpha. Records of its use in finance date back to the earliest statistics-based models in the 1960s.

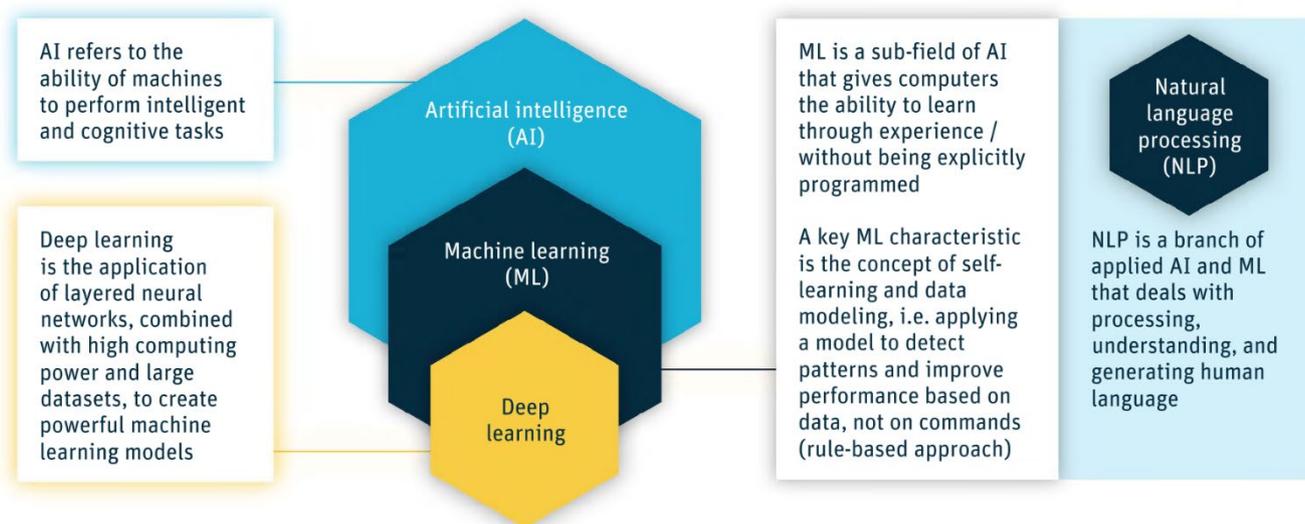
As the use of AI in the financial services industry gained steam, data-generation took new forms and new datasets, like ESG risk data, emerged. Used in tandem with traditional data like stock prices, ESG risk data is used as an overlay to inform and improve decision-making.

Decision-making utility can only be derived from relevant and precise datasets – on ESG risk or otherwise – but such datasets are not always what is produced and available to users. Ultimately, not all datasets are created equal.

In this article, we'll examine supervised machine learning and why we believe it produces the most superior datasets. In order to do this, let's define some common terms:

- **Artificial Intelligence (AI)** is the ability of machines to perform intelligent and cognitive tasks.
- **Machine Learning (ML)** is a sub-field of AI that gives computers the capacity to learn through experience without being explicitly programmed.
  - **Supervised machine learning** is a model of machine learning in which the algorithm learns on a human-labeled or annotated dataset.

- **Unsupervised machine learning**, by contrast, is a model of machine learning in which the algorithm is provided unlabeled data and tries to extract features or patterns on its own.
- **Natural Language Processing (NLP)** is a branch of applied AI and ML that deals with processing, understanding, and generating human language.



Supervised machine learning is found to yield up to eight times more strategic value than unsupervised machine learning.<sup>1</sup> This significant difference in strategic value is due to leveraging the best parts of artificial and human intelligence, while eradicating their respective shortcomings. Relying solely on artificial intelligence or unsupervised machine learning – say, a keyword search – produces large amounts of unrefined data very quickly. Relying solely on human intelligence and research produces small amounts of refined data very slowly.

Conversely, supervised machine learning quickly produces large amounts of refined data. The size and speed of AI enhances humans’ ability for parsing and utilizing its data, and humans can guide and refine the AI’s algorithms for more meaningful and accurate results – resulting in the significantly higher strategic value.

Currently, RepRisk is the only ESG research company leveraging supervised machine learning. Since 2007, RepRisk has produced the largest, high-quality annotated (human-labeled) dataset. This 14 year annotated data history, the longest consistent timeseries in the ESG industry, is a crucial element of supervised machine learning: the input of more than 500 years of human analyst work allows us to train our machine learning algorithms to be more accurate and effective in identifying ESG risks compared to an algorithm that learns based

<sup>1</sup> <https://www.gartner.com/en/newsroom/press-releases/2019-07-15-gartner-survey-reveals-leading-organizations-expect-t>, AI and Data Science in Trading - Digital Week 2020. AI Global. (2020). *Moving beyond the hype of AI & ML: Practical examples from other industries.* <https://www.aidatatradingdigital.com/landing/aidst-digital-week-on-demand>



on a non-annotated dataset. RepRisk's AI provides the capacity to screen more than 90,000 data sources in 20 languages for ESG risks every day, while over 85 highly-trained analysts curate and analyze the incidents according to a rules-based methodology for relevant and actionable insights.

The strategic value of alternative datasets, in particular ESG data, in the financial sector is becoming increasingly visible. As only relevant data has decision-making utility, supervised machine learning is emerging as the most effective mechanism to generate strategic value for businesses. RepRisk and its supervised machine learning-based approach remains on the cutting edge of data science to provide the highest-quality ESG risk data. See how our unrivaled dataset can be put to work for you at [reprisk.com](https://reprisk.com).