Forget Chips, AI Firms Need Higher Quality Data to Win

To build real artificial intelligence (AI), high-quality training data is arguably the single most important ingredient. As computing power becomes more of a commodity with each passing day, AI-focused firms, like Anthropic, Microsoft (MSFT), Alphabet (GOOG), Broadcom (AVGO), Oracle (ORCL), and Amazon (AMZN), are shifting their focus from building datacenters and buying as many Nvidia (NVDA) chips as possible. Instead, they are looking for truly reliable, high-quality datasets to power their AI's ability to learn, reason, and perform with anything close to human expertise.

Garbage in, garbage out. The same old saying from my early days coding in BASIC applies just as well to AI today. AI models require well-structured, relevant, and accurate data to achieve meaningful results, especially in complex fields like finance and investing.

As the critical role of high-quality training data becomes more widely recognized, AI firms are increasingly prioritizing investments in data acquisition and curation. However, these kinds of high-quality datasets are rare and hard to find. So much so, that we see companies taking rather extreme measure to get or create them.

For example, take Anthropic's approach to training its AI, Claude. Initially, Anthropic relied on broad internet data. Eventually, the company transitioned to more structured and vetted sources, such as newspapers, and most recently, books. Anthropic reportedly purchased millions of print books, digitized them, and used that content as a more refined and consistent training foundation for its models.

One critical question remains though - how do we imbue machines with the experience and expertise not found in books so they can perform tasks with comparable skill to human experts?

The answer lies in two foundational elements of data science: **taxonomy** and **ontology**.

In the context of an investing AI agent, taxonomy refers to the rigorous organization and classification of financial data. Excelling in this process for the investing business requires going beyond surface-level metrics and building advanced financial models that incorporate detailed, often-overlooked disclosures, such as those found in footnotes and MD&A sections of filings, as we've pointed out many times.

These sections frequently contain crucial information about subtle and nuanced accounting anomalies, off-balance-sheet items, and forward-looking risks that are essential to understanding a company's true financial condition.

Ontology refers to the structure that allows one to derive meaning from the data. It is the process of mapping the taxonomized data into relationships and patterns that can inform decision-making. For the investing business, the ideal ontology is one that can produce idiosyncratic alpha.

For ontologies to work, they need accurate and consistent data. No ontology, nor any model for that matter, can produce logical and meaningful signals if the underlying data is poor.

Accordingly, we have gone to great effort to build technology that gathers proven-superior data from financial filings to train and power our AI-Agent, the Robo-Analyst.

We have unrivaled experience and success in endowing machines with the subject matter expertise needed to perform like human experts. Want proof. First, this paper from the Harvard Business School and MIT Sloan professors empirically proves the idiosyncratic alpha in our proprietary measure of Core Earnings.

Second, the indices based on our research and managed by Bloomberg, strongly outperformed the S&P 500 in 1H25 and over the last five years. See Figures 1, 2, and 3.

1. Bloomberg New Constructs Ratings VA-1 Index (ticker: BNCVA1T:IND)
2. Bloomberg New Constructs Core Earnings Leaders Index (ticker: BCORET:IND)
3. Bloomberg New Constructs 500 Index (ticker: B500NCT:IND)

The "Very Attractive Stocks" Index beat the S&P 500 by over 62% over the last five years. Bloomberg's official name for the index is Bloomberg New Constructs Ratings VA-1Index (ticker: BNCVA1T:IND). Figure 1 shows it was up 158% while the S&P 500 was up 96%.

Figure 1: Very Attractive-Rated Stocks Index Strongly Outperforms the S&P 500 Over the Last 5 Years

Sources: Bloomberg as of July 18, 2025
Note: Past performance is no guarantee of future results.

The Bloomberg New Constructs Core Earnings Leaders Index beat the S&P 500 by 45% over the last five years. Our Index (ticker: BCORET:IND) was up 140% while the S&P 500 was up 96%. See Figure 2 for details.

Figure 2: Bloomberg New Constructs Core Earnings Leaders Index Outperforms the S&P 500 Over 5 Yrs

Sources: Bloomberg as of July 18, 2025
Note: Past performance is no guarantee of future results.

Our "Core-Earnings Weighted S&P 500" Index beat the S&P 500 by over 31% over the last five years. Bloomberg's official name for the index is Bloomberg New Constructs 500 Total Return Index (ticker: B500NCT:IND). Figure 3 shows it was up 128% while the S&P 500 was up 96%.

Figure 3: Bloomberg New Constructs 500 Index Strongly Outperforms the S&P 500 Over the Last 5 Yrs

Sources: Bloomberg as of July 18, 2025
Note: Past performance is no guarantee of future results.

Note that these indices are not available to the public. The only way to build strategies that achieve this kind of outperformance based on superior fundamental data is to be a New Constructs member.

Want to leverage our proven superior Robo-Analyst that can generate alpha in any market?

Learn more about our memberships here.

This article was originally published on July 21, 2025.

Disclosure: David Trainer, Kyle Guske II, and Hakan Salt receive no compensation to write about any specific stock, style, or theme.

Questions on this report or others? Join our online community and connect with us directly.

It's Official: We Deliver the Best Fundamental Data in the World

Many firms claim their research is superior, but none of them can prove it with independent studies from highly-respected institutions as we can. Three different papers from both the public and private sectors show:

1. The stock market is missing footnotes – and only we have that critical data.
2. Legacy fundamental datasets suffer from significant inaccuracies, omissions, and biases.
3. Our proprietary drives novel alpha. Our measures of Core Earnings and Earnings Distortion materially improve stock picking and forecasting of profits.

Best Fundamental Data in the World

In The Journal of Financial Economics, a top peer-reviewed journal, Core Earnings: New Data & Evidence proves our Robo-Analyst technology overcomes material shortcomings in legacy firms' data collection processes to provide superior fundamental data, earnings models, and research. More details.

Key quotes from the paper:

- "[New Constructs'] *Total Adjustments* differs significantly from the items identified and excluded from Compustat's adjusted earnings measures. For example… 50% to 70% of the variation in *Total Adjustments* is not explained by *S&P Global's (SPGI) Adjustments* individually." – pp. 14, 1st para.

- "A final source of differences [between New Constructs' and S&P Global's data] is due to data collection oversights…we identified cases where Compustat did not collect information relating to firms' income that is useful in assessing core earnings." – pp. 16, 2nd para.

Superior Models

Ernst & Young features the superiority of our ROIC, NOPAT and Invested Capital research to Capital IQ & Bloomberg's in Getting ROIC Right. See the Appendix for direct comparison details.

Key quotes from the paper:

- "…an accurate calculation of ROIC requires more diligence than often occurs in some of the common, off-the-shelf ROIC calculations. Only by scouring the footnotes and the MD&A [as New Constructs does] can investors get an accurate calculation of ROIC." – pp. 8, 5th para.

- "The majority of the difference…comes from New Constructs' machine learning approach, which leverages technology to calculate ROIC by applying accounting adjustments that may be buried deeply in the footnotes across thousands of companies." – pp. 4, 2nd para.

Superior Stock Ratings

Robo-Analysts' stock ratings outperform those from human analysts as shown in this paper from Harvard Business School. Bloomberg features the paper here.

Key quotes from the paper:

- "the portfolios formed following the buy recommendations of Robo-Analysts earn abnormal returns that are statistically and economically significant." – pp. 6, 3rd para.

- "Our results ultimately suggest that Robo-Analysts are a valuable, alternative information intermediary to traditional sell-side analysts." – pp. 20, 3rd para.

Our mission is to provide the best fundamental analysis of public and private businesses in the world and make it affordable for all investors, not just Wall Street insiders.

We believe every investor deserves to know the whole truth about the profitability and valuation of any company they consider for investment. More details on our cutting-edge technology and how we use it are here.

DISCLOSURES

New Constructs®, LLC (together with any subsidiaries and/or affiliates, "New Constructs") is an independent organization with no management ties to the companies it covers. None of the members of New Constructs' management team or the management team of any New Constructs' affiliate holds a seat on the Board of Directors of any of the companies New Constructs covers. New Constructs does not perform any investment or merchant banking functions and does not operate a trading desk.

New Constructs' Stock Ownership Policy prevents any of its employees or managers from engaging in Insider Trading and restricts any trading whereby an employee may exploit inside information regarding our stock research. In addition, employees and managers of the company are bound by a code of ethics that restricts them from purchasing or selling a security that they know or should have known was under consideration for inclusion in a New Constructs report nor may they purchase or sell a security for the first two days after New Constructs issues a report on that security.

DISCLAIMERS

The information and opinions presented in this report are provided to you for information purposes only and are not to be used or considered as an offer or solicitation of an offer to buy or sell securities or other financial instruments. New Constructs has not taken any steps to ensure that the securities referred to in this report are suitable for any particular investor and nothing in this report constitutes investment, legal, accounting or tax advice. This report includes general information that does not take into account your individual circumstance, financial situation or needs, nor does it represent a personal recommendation to you. The investments or services contained or referred to in this report may not be suitable for you and it is recommended that you consult an independent investment advisor if you are in doubt about any such investments or investment services.

Information and opinions presented in this report have been obtained or derived from sources believed by New Constructs to be reliable, but New Constructs makes no representation as to their accuracy, authority, usefulness, reliability, timeliness or completeness. New Constructs accepts no liability for loss arising from the use of the information presented in this report, and New Constructs makes no warranty as to results that may be obtained from the information presented in this report. Past performance should not be taken as an indication or guarantee of future performance, and no representation or warranty, express or implied, is made regarding future performance. Information and opinions contained in this report reflect a judgment at its original date of publication by New Constructs and are subject to change without notice. New Constructs may have issued, and may in the future issue, other reports that are inconsistent with, and reach different conclusions from, the information presented in this report. Those reports reflect the different assumptions, views and analytical methods of the analysts who prepared them and New Constructs is under no obligation to insure that such other reports are brought to the attention of any recipient of this report.

New Constructs' reports are intended for distribution to its professional and institutional investor customers. Recipients who are not professionals or institutional investor customers of New Constructs should seek the advice of their independent financial advisor prior to making any investment decision or for any necessary explanation of its contents.

This report is not directed to, or intended for distribution to or use by, any person or entity who is a citizen or resident of or located in any locality, state, country or jurisdiction where such distribution, publication, availability or use would be contrary to law or regulation or which would be subject New Constructs to any registration or licensing requirement within such jurisdiction.

This report may provide the addresses of websites. Except to the extent to which the report refers to New Constructs own website material, New Constructs has not reviewed the linked site and takes no responsibility for the content therein. Such address or hyperlink (including addresses or hyperlinks to New Constructs own website material) is provided solely for your convenience and the information and content of the linked site do not in any way form part of this report. Accessing such websites or following such hyperlink through this report shall be at your own risk.

All material in this report is the property of, and under copyright, of New Constructs. None of the contents, nor any copy of it, may be altered in any way, copied, or distributed or transmitted to any other party without the prior express written consent of New Constructs. All trademarks, service marks and logos used in this report are trademarks or service marks or registered trademarks or service marks of New Constructs.

Copyright New Constructs, LLC 2003 through the present date. All rights reserved.