Machine Learning & AI for Smarter Investing

New Constructs®

Table of Contents

New Constructs®

Diligence | Independence | Performance

www.newconstructs.com

Part 1: Cutting Through the Smoke and Mirrors of AI on Wall Street

Artificial intelligence made lots of headlines in 2017. Alphabet (GOOGL) developed software that defeated the defending world champion in Go, then a few months later developed a new version that defeated the prior version 100 games to none.

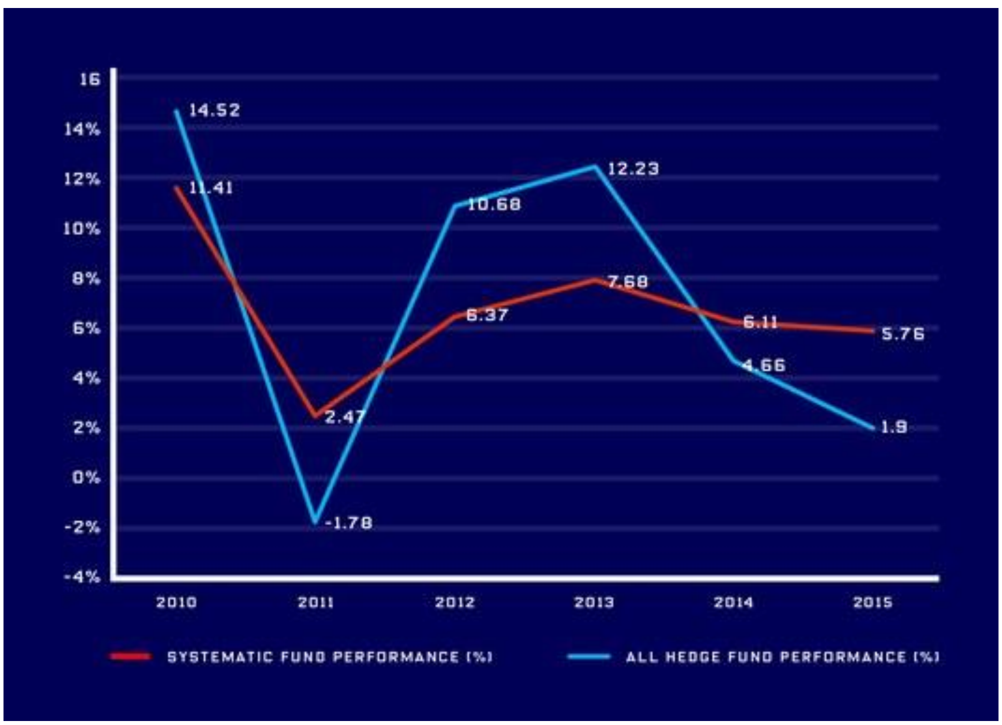
These developments have spurred predictions that "AI Will Invade Every Corner of Wall Street." Prognosticators see a world in which computers completely replace human investors.

"If computing power and data generation keep growing at the current rate, then machine learning could be involved in 99 percent of investment management in 25 years," Luke Ellis, CEO of fund management company Man Group, PLC, told Bloomberg.

Despite this optimism, advances in artificial intelligence have not yet translated to superior returns. According to *Wired*, quant funds over the past few years have, on average, failed to outperform hedge funds (which have themselves failed to outperform the market).

Most people do not understand that AI, especially the AI used in finance today, lacks the application of deep subject matter expertise[1] to create the clean data and relationships that are the foundation of any successful investment strategy or AI. Winning games is one thing, but the real world is not a game that follows immutable rules in a strictly defined space. In the real world, humans change the rules, break the rules, or the rules don't even exist. Current AI is nowhere near navigating real world situations without a great deal of human intervention.

Figure 1: AI Is Overhyped and Misunderstood: Systematic Funds Underperform



Sources: Preqin/Wired

[1] Harvard Business School features the powerful impact of our research automation technology in the case New Constructs: Disrupting Fundamental Analysis with Robo-Analysts.

Finding the Talent(s)

One of the biggest problems with AI today is lack of interest or ability of those with adequate subject matter expertise to communicate with the programmers building the AI. The programmers don't understand the data they're feeding into their AI, and the analysts lack the understanding of the technology to communicate what programmers need to know to understand the source data and interpret the results.

This disconnect creates a number of well-publicized issues for the application of AI in finance and investing:

- Most AI firms end up spending the bulk of their resources on data management and data scrubbing rather than technology.
- Machines often discover spurious correlations that don't work, or only worked in the past but aren't applicable in the future.
- Many AI systems turn into "black boxes" that spit out investment recommendations with no explicable basis or strategy. If the AI cannot articulate to humans how it "thinks," then how can investors trust it with significant sums of money?

Individuals with the skills and knowledge to bridge this divide are among the scarcest and most valuable people in finance. Nine out of 10 financial services firms have already started working on AI technologies, and they're all competing in this scarce labor pool.

As we wrote in "Big Banks Will Win the Fintech Revolution," the largest financial firms will be the biggest beneficiaries of technological advancements due to their scale and resources. Big banks can afford to pay the most for AI talent, and they have the biggest store of financial data to aid their new programmers.

A few banks are already making serious efforts to get the necessary talent. UBS (UBS) is on an AI hiring spree, while Morgan Stanley's (MS) programmers and financial advisors have worked together to build "Next Best Action", a platform that uses machine learning to aid its advisors in offering personalized advice to clients.

These efforts should eventually pay off in a big way, but for now they remain in their infancy. Financial institutions still have a long way to go before they can truly implement AI in an effective way.

The Big (Data) Problem with AI

The total amount of digital data in the world doubles every two years. As the volume of data grows exponentially, most of that data lacks the structures needed for machines to analyze it. As a result, AI projects, which are supposed to reduce the need for human labor, require countless man-hours to collect, scrub, and format data inputs.

Virtova founder, Sultan Meghji, told the Financial Revolutionists that many AI startups spend at least half their funding on data cleanup and management. Everyone wants to talk about teaching computers to think, but there's no short cut or substitute for curating the data sets that machines use to learn.

To train an AI, you need a training data set for it to learn from. Training data sets tend to be of two kinds. First, you have relatively small, accurate data sets that don't contain enough different kinds of examples to be effective. AI trained on these data sets become great at interpreting the training data, but they can't handle the variety and vagaries of the real world.

Other training sets are large but not very accurate. In these case, the AI gets to see lots of examples, sometimes with incorrect data, but it isn't being given clear and consistent instructions on how to respond. AI trained on these larger, inaccurate data sets often determine that there are few consistent things to be learned from the data and are capable of doing very little on their own.

For successful machine learning, training data sets need to be both accurate and widely representative. In other words, the training data needs to accurately represent what happens in as much of the real world as possible. How else can we expect the machine to learn anything consistently useful?

Herein is the AI challenge: machines can't learn without good training data sets and creating good training data sets requires more time than most realize from humans with deep subject matter expertise.

Most humans with the depth of subject matter expertise required to curate a good training data set are not interested in such mundane work. An alternative approach is to have lots of humans with limited subject-matter-expertise do the work, but this approach has been unsuccessful so far.

The Big (Data) Problems Are Worse in the Finance & Investing World

In theory, curating training data sets should be less challenging in finance. After all, financial data is structured in the form of financial statements in official filings with the SEC. However, any layman can quickly see that there is not as much structure (humans do not always follow the rules) as one might presume in these filings. Plus, the structure that does exist is not all that useful for AI. In fact, it can be actively harmful.

Imagine a computer that wants to compare the financials of Coca-Cola (KO) and Pepsi (PEP). As the computer reads through the financial statements, how is it supposed to know that “Equity Method Investments” for KO and “Investments in Noncontrolled Affiliates” for PEP are the same? What about “Retained Earnings” vs. “Reinvested Earnings.” Industry groups have been trying to create a standardized financial nomenclature for years to solve this very problem.

In theory, the development of XBRL would solve this problem. In practice, XBRL still contains too many errors and custom tags to allow for fully automated reading of financial filings. Even the smartest machines need extensive training from humans with deep subject-matter expertise to be able to understand financial filings.

Without this pairing of sophisticated technology and expert analysts, any AI effort in finance is doomed to failure. As the saying goes, “garbage in, garbage out.” Dumping a bunch of unstructured, unverified data into a computer and expecting it to deliver an investment strategy is like dumping the contents of your pantry into the oven and expecting it to bake a pie. It doesn’t matter how good the machine is, it can’t function without the right preparation.

The Problem of False Positives

Even if the financial data is structured and verified, it may not be useful to a machine, and AI will struggle to tell what data is useful and what is not. The large volume of available financial data means there will inevitably be a large number of apparent patterns that are actually the result of pure randomness. This phenomenon is known as “overfitting,” and it’s such a recognized issue that it gets its own lesson in Stanford’s online course on machine learning.

Overfitting is not just an AI problem. Humans have always struggled with seeing patterns where none truly exist (heuristics). At least, though, we can be conscious of this flaw and try to counteract it. Computers, for all their sophistication, cannot claim this same level of consciousness. When programmers design machines to find patterns, that’s what those machines are going to do.

As AI gets more complex, the problem of overfitting gets worse. Anthony Ledford, the chief data scientist at one of Man Group’s quant funds, recently told *The Wall Street Journal*:

“The more complicated your model the better it is at explaining the data you use for training and the less good it is about explaining the data in the future.”

Many quant funds today are simply mining patterns from past data and hoping those patterns persist into the future. In reality, most of those patterns were either the result of randomness or conditions that no longer exist.

Again, we see the need for the pairing of AI with human intelligence. Machines can process data and find patterns more quickly and efficiently than any human, but for now they lack the intelligence to audit those patterns and understand whether or not they can be used to predict future results.

AI as a Black Box

Of course, to audit the results of AI, humans need to be able to understand how that AI thinks. They need some level of insight into the processes the machine is using and the patterns it discovers.

Right now, most AI is not transparent enough for potential users to trust it. All too often, the AI algorithms are a black box that take in data and spit out results without any transparency into their underlying machinations.

In part, this problem is unavoidable if we want the machines to operate with the scale needed for them to be useful. The code that goes into AI is so complex that few individuals could ever fully understand its inner workings.

In fact, software doesn't even have to reach the complexity of AI to have these problems. Consider the unexpected acceleration problems that plagued the Toyota Camry about 10 years ago. So many programmers had worked on the engine control software that it turned into "spaghetti code," a mass of unintelligible and often contradictory code that no one understood and caused great harm.

If the software to support human control of a car's breaking and acceleration can become so complex, just imagine how much more confusing and susceptible to errors more sophisticated activities, like financial modeling, can be. One mistake in one line of code could alter the entire function of the system. The software wouldn't break, it would just be performing a different task than intended without anyone realizing until, perhaps, it's too late.

This problem is exacerbated by the divide between the people with adequate subject matter expertise in finance and the programmers. The finance experts don't understand how the software works, while the programmers don't understand how finance works.

Finance is far from the only sector to experience this problem. In "The Coming Software Apocalypse," *The Atlantic* detailed several examples of major failures that occurred because the coders didn't properly anticipate all the potential uses of their software. These failures were prolonged because the people using the code didn't have any idea how it worked.

As long as AI remains a black box, its utility will be limited. Eventually, the lack of transparency will lead to a significant and undetected failure. Even before that point, it will be difficult to get investors to commit significant money to a program they cannot trust.

The Way Forward

For all these challenges, AI will continue to expand its reach on Wall Street. There's no other way for financial firms to meet the dual mandate of reducing costs and improving their service. Technology is the only solution for analyzing the huge volumes of corporate financial data filed with the SEC every hour and meeting the Fiduciary Duty of Care.

The firms that understand this fact and take concrete steps to invest in technology will have a significant advantage over their competitors, which is why UBS and Morgan Stanley are among our top picks in the financial sector.

Part 2: Opening the Black Box: Why AI Needs to Be Transparent

Just as ancient man attributed natural phenomena to the gods without question, so too do many modern humans blindly trust in the power of AI. However, for AI to be a truly effective tool in finance, transparency about its inner workings should be a natural by-product of the integrity of its design and effectiveness.[2]

"I have bought this wonderful machine—a computer. Now I am rather an authority on gods, so I identified the machine—it seems to me to be an Old Testament god with a lot of rules and no mercy."

-Joseph Campbell, The Power of Myth

The idea of the computer as a god has existed almost as long as the computer itself. In 1954, science fiction writer Isaac Asimov turned the computer into a literal god in his short story, "The Last Question." In this story, generation after generation of humans beseech a galactic computer to stave off the end of the universe, but only after the universe dies does this computer begin the act of recreation with the words, "LET THERE BE LIGHT!"

Even working with the primitive computers of the 1950's, Asimov understood how complex the machines could become:

"Multivac (the computer) was self-adjusting and self-correcting. It had to be, for nothing human could adjust and correct it quickly enough or even adequately enough. So Adell and Lupov attended the monstrous giant only lightly and superficially, yet as well as any men could. They fed it data, adjusted questions to its needs and translated the answers that were issued."

Anyone working in AI today doubtless recognizes themselves in Adell and Lupov. They feed data in, try to make sense of the results, and do their best to maintain a black box of a system no one truly understands.

How We Got Here

The development of artificial intelligence as a black box may have been predictable, but it was far from inevitable. The lack of transparency came about due to the fact that the people developing the backbone of AI systems and the ones tasked with deploying those systems have little in common.

In fact, these disparate folks tend not to work under the same roof. While a large percentage of companies claim to be developing AI technologies, the truth is that most of them are just repurposing others' research. The bulk of AI programs are being developed in universities and by a handful of tech giants like Google (GOOGL), Microsoft (MSFT), and Facebook (FB).

The researchers at these institutions create AI algorithms for a variety of purposes. Developers at other companies discover that these algorithms can do 80% of what they need and hack together modifications to try to get the other 20%.

The end result is a program that no one fully understands. The initial researchers don't know how the AI is being used. The developers don't understand how the underlying code, that they get from some another party, works (or, more often, does not work). The actual users couldn't get a clear picture of how the AI makes decisions even if they tried – which they usually don't.

Another part of the problem is simply the newness of AI. Researchers have been so focused on teaching computers to "think" in a way that produces useful results that they haven't spared as much thought on making those thought processes transparent. Further, given the lack of success in getting AI to produce useful results, there are not a lot of inner workings to brag about. Why show off something that does not work so well?

Slowly but surely, the technological community is beginning to question the black box model. After all, AI exists because it's supposed to outstrip human intelligence at specific tasks. If the AI's thought process is

[2] Harvard Business School features the powerful impact of our research automation technology in the case New Constructs: Disrupting Fundamental Analysis with Robo-Analysts.

superior, shouldn't its explanation of that thought process be equally sophisticated? As philosopher and cognitive scientist Daniel Dennett told the MIT Technology Review:

"If it can't do better than us at explaining what it's doing, then don't trust it."

The Problems of the Black Box

Dennett identifies the most pressing reason why AI needs to become more transparent. As long as AI stays a black box, many will be hesitant to trust it. When it comes to handling people's money, this lack of trust becomes even more pronounced.

The common rebuttal to this concern has been that the superior returns of AI-driven investment decisions will win skeptics over. Putting aside the fact that AI hedge funds haven't yet succeeded in delivering outperformance, there's reason to believe that returns alone won't be enough to attract investor capital.

Research shows that the best performing asset managers don't necessarily attract the most new capital. The most successful funds at attracting new capital do so by building investor trust through transparency and investor education. Many investors prefer solid returns they understand over great returns they don't.

These investors are right not to trust black boxes. Even if AI succeeds in identifying unseen patterns to deliver superior returns, that success will be only temporary without transparency. If no one understands how the machine works, they won't be able to figure out when it stops working correctly, much less how to fix it.

Right now, many companies that implement AI have to do so through a process of trial and error. If the machine fails to correctly process a data set, they just have to keep (often blindly) tweaking different parameters up and down until it spits out the right results.

Trial and error might work in the "move fast and break things" world of Silicon Valley, but it's not going to fly when investors' money is on the line.

How to Open the Black Box

As AI becomes more ingrained in every part of our lives, people are making a concerted effort to fix the technology's transparency problem. The AI Now Institute recently issued a set of recommendations for the AI industry: goal number 1 was that core public agencies should no longer use black box AI systems.

On the finance side, the EU's General Data Protection Regulation could impose punishments on companies that rely on black box algorithms. While the Fiduciary Rule doesn't touch on AI specifically, you have to believe that black box technologies won't be enough to fulfill the Duty of Care. What regulator would accept "the computer told me to do it" as the rationale for an investment decision?

In order to fulfill the goal of transparency, researchers have put a lot of effort into teaching AI to explain itself. Researchers at MIT successfully built a neural network with two different modules that could successfully make simple predictions and highlight its reasoning. Developments such as these prove that the black box can be opened.

In order to end the black box, financial companies working on AI should take a few concrete steps:

1. Take a holistic approach to development where those with tech expertise and those with finance expertise work together on every stage of the project.
2. Make explicability a key goal. An AI that can't explain its reasoning is no true AI.
3. Commit to auditability. As much as possible, make the data inputs and outputs accessible so that others can verify the machine's "thought" processes.

These steps seem obvious. Point 1 is clearly the best strategy for developing any technology, and 2 and 3 are not only good in principle, they're good for business. If you've done the work to build great technology, why wouldn't you want to show clients all the work you've done? It's easier to build trust in products clients can see and understand.

We've seen progress on point 1. Certain financial firms are making an effort to bring more tech knowledge in-house, as banks like JPMorgan Chase (JPM) are going head-to-head with tech giants to hire AI researchers and developers.

The second point has been more of a mixed bag. Credit card companies like Capital One (COF) have dedicated research teams trying to make their computer techniques more explainable. On the other hand, many asset managers still seem content to rely on black box technology.

The final point can be the most difficult, but it's also incredibly important. Financial firms often use the sensitivity of their data as an excuse to keep it hidden even when they could potentially anonymize and share it. Without the availability of this data, it becomes much harder for investors to trust in AI. No matter how sophisticated the technology is, it will be useless without the right data. Garbage in, garbage out.

Speaking from Experience in Building a Transparent Box

In the development of our machine learning and AI technology, we've found that commitment to a holistic approach, explicability and auditability pays many dividends. It boosts the trust of our clients in our work (e.g. our click-through filings service), and augments our ability to write better code and improve the AI.

On the holistic approach: since our inception, we've rallied around the benefits of combining different kinds of expertise. We eschew the idea of siloed teams. Empirical evidence proves the benefits of working in teams built on different experiences and skill sets. Open communication between analysts and programmers allows us to anticipate and address problems in advance while also reviewing our work with a critical eye. We see great competitive advantage in our ability to combine technological and finance expertise.

On explicability: by clearly documenting our code and using consistent formatting that everyone on the team understands, we're able to work efficiently with a larger team. We avoid any reliance on one person going down a rabbit hole that no one else understands. Explicability keeps all our programmers on the same page and provides excellent training material for bringing new programmers up to speed. It also ensures there is one version of the truth about how our code works.

In addition, all project requests from our financial analysts are required to conform to specific standards so that both parties know exactly what is expected of the other. We derived these standards over 15 years of successes and failures in machine learning and AI development. One of the most important benefits of these standards is the improved transparency and communication they create for the different minds working on projects. Plus, they ensure there is one version of the truth about how the teams will execute the project and test its success.

On transparency: having a better understanding of how code changes affect output makes communicating about how code does or does not work easier. When the team is able to directly and discretely measure the output or impact of code changes, then it can directly measure the efficacy of the code and communicate how it may need to change. Transparency enables everyone to be more aware and, therefore, smarter about the process and the results. Running all of our data through financial models that we publish to clients unifies the firm around a discrete set of outputs that we can manage and measure.

Transparency boosts external communications as well. We show the source filing data and all the calculations our Robo-Analyst uses to build our models and determine investment ratings. We want clients to understand how much work our machines do and be able to verify thought processes. There's no reason for us (or any company) to hide how much work, planning and sophistication goes into our technology.

Part 3: AI Has a Big (Data) Problem

The total size of all global data hit 20 zettabytes in 2017. For 99% of people, that number probably means nothing, so picture this: if every 64-gigabyte iPhone were a brick, we could build 80 Great Walls of China with the iPhones needed to store all the world's data.

We are awash in an ocean of data that grows bigger by the second. And it's a complete and utter mess.

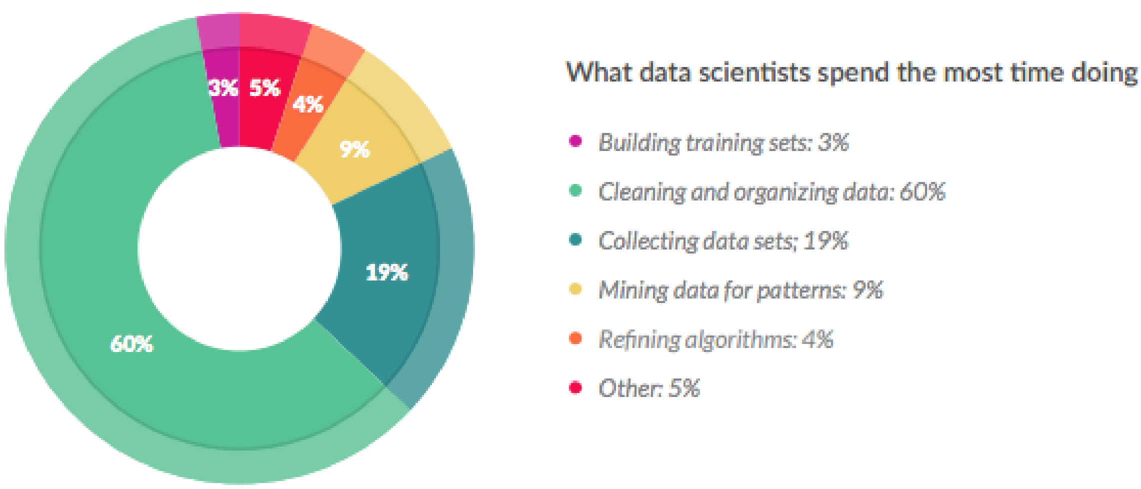
90% of web data is unstructured, meaning it's in a format that cannot be easily searched and understood by machines. Poor data quality costs the US economy $3.1 trillion a year according to IBM (IBM). We have become a society that is excellent at producing, storing and sharing data, but we're lousy at making it useful.

Poor data quality represents the single largest hurdle for developing useful artificial intelligence. It doesn't matter how "smart" machines become if they're fed data that is inaccurate or incomprehensible.

The Size of the Data Management Problem

Poor data quality is a familiar problem for those who analyze data for a living. A recent survey found that 60% of data scientists devote the majority of their time to cleaning and organizing data, as shown in Figure 2.

Figure 2: Cleaning Data Takes the Most Time



Sources: CrowdFlower

Comparatively, just 9% of data scientists devote the bulk of their time to mining data for patterns. Cleaning and organizing data has become such a big task that it leaves precious little time for analysis.

When people predict that AI will make human workers obsolete anytime in the near future, they are ignoring the data quality problem. AI and machine learning may be able to replace those 9% of data scientists who are mining data for patterns, but it will still need the 80% working on collecting, cleaning and organizing data.

More importantly, data scientists need to re-orient their thinking around data quality. More time needs to be spent upfront and on collecting data in a high-integrity manner rather than retroactively "cleaning" it. We are not sure that it is possible to retroactively clean data enough to meet the needs of useful AI. If you cannot validate the data back to its source, then how do you know it is clean? And, if you are going back to the source to validate, then you might as well collect it from the source.[3]

[3] Harvard Business School features the powerful impact of our research automation technology in the case New Constructs: Disrupting Fundamental Analysis with Robo-Analysts.

Gradually, leaders in AI seem to be understanding that their job is as much, if not more, about data management as it is about new technologies. Facebook (FB) just hired Jérôme Pesenti, the former head of IBM's Big Data group, to run its AI efforts. Pesenti replaces Yann LeCun, who will now focus on his core expertise of research.

While research into deep learning and neural networks makes the most headlines, Facebook understands that data management is crucial to delivering bottom-line business improvements.

AI Has a Long Way to Go

Once you understand the limitations imposed by poor data quality, the challenges facing AI become much clearer. Successes like the Go-playing computer that beat the world champion are misleading because they are enabled by completely structured and easily interpreted data points. Accordingly, the subset of tasks that AI can perform effectively remains small, according to LeCun:

"In particular areas machines have superhuman performance, but in terms of general intelligence we're not even close to a rat."

Other researchers have faced similar hurdles. Machines can handle finite and discrete data points well, but even a minor degree of ambiguity can trip them up.

"We are still a long way from computers being able to read and comprehend general text in the same way that humans can."

The last quote comes from Microsoft (MSFT) CTO, Kevin Scott, in a LinkedIn post celebrating the development of an AI that could read and answer questions about Wikipedia pages at the level of an average human. Despite its success, the machine struggled when asked to go beyond simple facts and make intuitive, but still logical, leaps. Contextual clues that a human would understand easily are still incomprehensible to machines.

This problem is even more pronounced when it comes to using machines to read financial filings. One of the biggest issues people face when trying to use natural language processing on financial filings is that the language in these documents is far from natural. If machines can get tripped up by Wikipedia, imagine how they respond to the jargon and legalese that fill your average 10-K.

In the financial world, as in the rest of the economy, AI has had its biggest successes working with data that is structured and standardized. JPMorgan Chase (JPM) has automated hundreds of thousands of hours of work annually by developing machine learning tools to read commercial-loan agreements. These contracts are standardized, which means the machine only has to navigate minor differences.

While machines are replacing humans in these rote tasks, they struggle to make the leap to more sophisticated analysis. Rather than try to tackle the complex task of reading and analyzing financial filings, most applications of AI in the investing business focus on mining patterns out of trading data, alternative data sets, or sentiment and other non-financial indicators. As we discussed in our first piece in this series, these efforts have not yet led to superior returns. It's been easier to apply existing AI and machine learning tools to new data sets than it has been to teach AI to analyze old data sets, especially data directly from SEC filings.

Structuring Financial Data: Not as Easy as Most Think

In theory, financial data in filings would be more structured and standardized, or we could make it that way easily. We have centralized bodies (FASB, SEC) that govern financial reporting standards. Public companies employ teams of accountants and lawyers to conform to these standards.

In reality, the data remains highly unstructured and variable, and we expect that it will only get worse. The most prominent effort to make financial data machine readable, XBRL, remains riddled with errors 10 years after its initial deployment. While companies are required to submit XBRL filings, they're not required to verify them, and only 8% of companies carry out voluntary audits. Until XBRL is strictly enforced by the SEC, it does not stand a chance at being reliable.

Without SEC enforcement, most companies will continue to under invest in the resources necessary to get their XBRL filings right. They will either leave the task in the hands of accountants that lack the technical expertise to do the job, or they outsource the job to a third-party that doesn't fully understand the company's financials.

Again, we come back to the disconnect between people with technical expertise and those with subject matter expertise in finance.

As long as XBRL and other efforts to structure financial data are treated as curiosities that companies can safely ignore, it will be almost impossible to make meaningful progress on this front.

Structuring Data: More About Team Than Technology

As long as financial data remains unstructured, existing machine learning tools cannot process it effectively. Meanwhile, the cost of employing the highly-trained analysts needed to manually structure data remains prohibitive.

Our solution is to leverage our deep financial expertise into software that enables highly-trained analysts to collect and structure data with unrivaled efficiency. In essence, we arm human subject matter experts (SMEs) with technology at every step of the data collection and modeling process. For example:

- Analyst and programmers work together to build the AI. Analysts and programmers anticipate and address problems from multiple perspectives from the outset. Clear communication between financial and technical experts is critical to building machines that work. Anticipating potential problems at the start, along with frequent iteration and joint and rigorous testing, helps build machines that robustly do something small. One step at a time, we teach machines to perform discrete tasks perfectly. However small that step may be, each step means less work for human SMEs. We don't send programmers or data scientists to analyze the data in isolation.
- During data collection, our process leverages a multitude of sophisticated algorithms to validate the data points collected by the machine so humans can transcend most of the banal work. We use our big data experience and financial expertise to automatically identify data that's potentially wrong – from values that are too big or too small, to data points that show up in the wrong places, to data relationships that don't make financial sense. Analysts only have to focus on issues the machines have not already mastered. We also track how analysts address each issue so that if it recurs, the machines can handle it automatically.
- Our data collection process includes sophisticated corporate performance and valuation modeling of the data that produces highly-respected investment ratings and research. Our financial expertise enables us to create quality assurance algorithms that flag modeled results that are unusual or have been linked to errors in the past. This process adds significant integrity to our data collection process compared to traditional data collection processes by humans who are not experts in accounting or finance or do not have a model to help them analyze the impact of the data they collect.
- To teach a machine well, we need to think like machines. Accordingly, every data point in the 120,000 filings parsed into the machine by human SMEs is tagged with 10+ pieces of unique identifying information. These tags include data value, the associated text, the location in the filing, and many other features that are taken for granted by most humans but provide critical context for the machine.
- The scale and efficiency of our process has a virtuous effect on our data validation processes. The more models we can build, the more potential data anomalies or errors we can find and feed back into the machine. The more we do, the more we can teach the machine, and, in turn, rely on it to do more. This approach gives us significant advantage over systems or analysts who can only view a few models at a time.

Working with machines presents many new challenges to our society. It is not something we've done before and, not surprisingly, we have a lot to learn, and so do the machines. One thing we know for sure is that people that are best at teaching machines will have the best machines, and the people with the best machines will have the upper hand.

Part 4: Working with Intelligent Machines

Discussions of artificial intelligence usually bring up one of three scenarios in people's minds:

1. **The Terminator Scenario:** AI becomes sentient and overthrows humanity in a violent revolution. Machines are dangerous and our enemy.
2. **The Jetsons Scenario:** AI exists in a state of total subservience to humans, like Rosie the robot maid in the Jetsons. Machines are our servants
3. **The Automation Scenario:** AI slowly replaces human workers until there are no jobs left. Society experiences massive upheaval as we adjust to a post-work society. Machines are a replacement for human labor.

Machines will destroy us, serve us, or replace us. Those are the three dominant ways society imagines humanity relating to artificial intelligence.

The Terminator and Jetsons scenarios have been the most widely discussed in fiction and entertainment (hence their names), but it's the automation scenario that now dominates contemporary thinking around AI. Every day brings another breathless prediction that robots will take all our jobs. One of the most prominent advocates of this position is Tesla (TSLA) CEO Elon Musk, who told the National Governors Association:

"There certainly will be job disruption. Because what's going to happen is robots will be able to do everything better than us. ... I mean all of us."

When Musk says "robots will be able to do everything better than us," he makes two key errors. He overestimates robots, and he underestimates humans. AI will not replace humans, and it won't destroy us or be totally subservient to us either. Instead humans and AI will work together[4], combining each other's strengths and compensating for each other's weaknesses to create jobs and achieve results that are presently unimaginable.

The Limitations of AI

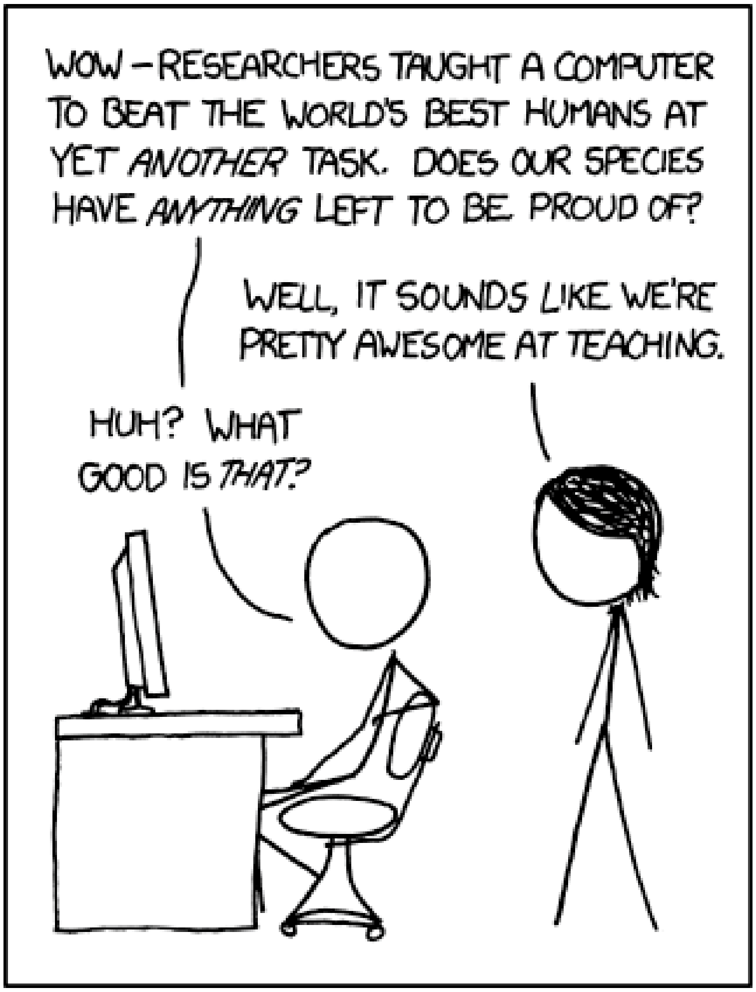
As we discussed in "Cutting Through the Smoke and Mirrors of AI on Wall Street" (and elaborated on in the second and third articles of this series), AI has a long way to go before it can compete with human intelligence. Machines may be able to defeat humans in games like Chess and Go, but they can't compete with humans in areas such as logical intuition, much less creativity and innovation. Machines are savants: incredibly skilled at specific tasks but limited in their overall cognition.

The idea that AI will replace all human workers ignores the complexity and variability of most jobs. Sure, you can train a machine to do x, but most jobs require people to seamlessly juggle tasks x, y, and z, often at the same time and in an open environment that is not immutably structured (i.e. a chess board).

Even if AI does advance to match humanity's level of general intelligence, it will still need to learn how to apply that intelligence. Smart machines need even smarter teachers, and humans are the best teachers around.

[4] Harvard Business School features the powerful impact of our research automation technology in the case New Constructs: Disrupting Fundamental Analysis with Robo-Analysts.

Figure 3: Teaching Machines



Sources: xkcd

Machines are undoubtedly superior to humans at a large number of confined or routine tasks, but they will never "be able to do everything better than us."

Machines Can Empower Us & Make Us More Human

No one understands the pain of being surpassed by a machine better than Gary Kasparov. When the reigning world chess champion was defeated by IBM (IBM) supercomputer Deep Blue in 1997, Newsweek described it as "The Brain's Last Stand."

Today, Kasparov sees his defeat differently. He believes the advancement of artificial intelligence represents a boon to humanity, even if machines do cause some disruption to the labor market. The more that AI replaces routine jobs, the more it frees up humans to dream up jobs that could never have existed in the past, as he wrote in an essay last year in the *Wall Street Journal*.

> *"Machines that replace physical labor have allowed us to focus more on what makes us human: our minds. Intelligent machines will continue that process, taking over the more menial aspects of cognition and elevating our mental lives toward creativity, curiosity, beauty and joy. These are what truly make us human, not any particular activity or skill like swinging a hammer—or even playing chess."*

Think of the jobs that exist today that would have been unimaginable 20 years ago. Now think of the jobs that will exist 20 years from now. Just like it would be hard to explain the concept of a social media manager to someone from 1998, so too is it difficult to comprehend the jobs that will exist in 2038. As Box (BOX) CEO Aaron Levie explained on Twitter:

"AI can seem dystopian because it's easier to describe existing jobs disappearing than to imagine industries that never existed appearing."

Humans are wired to fear the unknown. The loss of jobs that already exist to machines frightens us more than the potential jobs those machines will create, even though the new jobs will make use of our talents in ways that are more rewarding and productive.

How We Can Work with Machines

The cycle above—machines taking over human jobs, which then frees those humans to create new jobs—has existed for centuries. Agricultural advancements allowed farmers to move to cities and become artisans and merchants. The Industrial Revolution displaced those artisans and created jobs building railroads and steamships. Throughout history we have made technological advancements that replaced some jobs and created new ones.

What's different about AI is our relationship to the technology. AI is not a tool where the human is in complete control. AI gets some autonomy. It acts in ways that it's not explicitly told to and makes recommendations that humans might not anticipate or fully be able to understand. People don't just need to learn how to use AI, they need to learn how to work cooperatively with it.

Many large financial firms have already shown an understanding that AI is best used as a complement to human labor, not a replacement for it. When BlackRock (BLK) announced last year that it would rely more heavily on algorithms to pick stocks for its funds, much of the news coverage interpreted the move as machines replacing traditional fund managers.

BlackRock, however, insists that's not the case. Cofounder Rob Kapito outlined the company's view on AI at the Barclays New Frontier Conference last November:

"It's not going to replace humans. I believe it will be human and machines."

The numbers bear Kapito's statement out. Even though the company's shift to more algorithmic funds led to 36 employees leaving, the firm has actually added nearly 700 employees (5% increase) over the past year. It's not just tech talent either; BlackRock has placed an emphasis on hiring liberal arts majors. As technology changes the finance industry, employers need to start looking for different skillsets.

Research shows that cooperation between humans and machines thrives when humans leverage our uniquely human skills: intuition, pattern recognition, and reading implicit signals. Relationships thrive when both parties have a clear and early understanding of what to expect from each other.

AI has the potential to make us more human, not less, when developed with these goals:

1. **Practice Transparency:** In a previous article, we explained the challenges for humans to work with black boxes. Programmers and subject matter experts (SMEs) need to work together to ensure the abstract goals for the machine are properly and clearly expressed in the code. In practice, achieving this goal requires disciplined adherence to clear, documented processes for communication and code writing that both the programmers and SMEs understand.
2. **Understand Strengths & Weaknesses:** Machines and humans are most effective together when machines perform tasks that play to their strengths (sorting through large amounts of data, statistical analysis, identifying patterns, etc.) while humans focus on their strengths (putting data in context, interpreting results, understanding priorities, etc.). Get the little things right before adding complexity. Don't expect too much from your machines or your humans.
3. **Integrate Diverse Minds:** If there's one theme we keep coming back to in this series, it's the need to integrate technological expertise and with subject matter expertise. The companies that are succeeding with AI enable programmers and SMEs to communicate effectively across all levels of the organization.

These steps can help companies reframe AI as way to make humans more productive rather than as a way to replace them. In the words of famed investor Paul Tudor Jones:

"No man is better than a machine, and no machine is better than a man with a machine."

Part 5: How AI Can Help Advisors Grow and Keep Assets

Financial advisors face unprecedented challenges today. On the one hand, fee compression has made it difficult to maintain profits. On the other, the Fiduciary Rule and competition from robo-advisory technology requires increasingly higher levels of service and costs.

We see technology as one of the only true solutions to these challenges. Specifically, artificial intelligence empowers research automation[5] that gives advisors the diligence and research they need to impress new and existing clients and while avoiding regulatory scrutiny.

Embrace the Change: AI Is an Opportunity, Not a Threat

Before long, using AI will no longer be optional for advisors. In the coming decades, $30 trillion of wealth will be transferred to millennial investors that expect their advisors to be tech-savvy. The CFA exam plans to require knowledge about AI beginning in 2019.

For many advisors, the inevitable rise of AI raises two fears:

1. AI will replace human advisors, and all investors will be served by robo-advisors.
2. If human advisors do survive, they will need coding skills that are difficult to learn mid-career.

While these fears are a natural reaction to change, we think they are unfounded. Investors young and old want a human touch and personalized level of service that robo-advisors can't deliver. They need a human they can rely on in a volatile market, not a robo-advisor that crashes when they try to check their accounts.

As we discussed in "AI Has a Big (Data) Problem", machines can excel at specific tasks, but they're nowhere near the level of sentience needed to deliver personalized service to a wide array of clients.

Nor will financial advisors need to become expert coders to work with AI. According to Stephen Horan, the managing director for credentialing for the CFA Institute, the technological hurdle is not nearly that high:

"Candidates will not be expected to code computer programs, but rather distinguish between structured and unstructured data analytic methods as well as identify characteristics of robust investment algorithms."

Working with AI doesn't mean you need a Ph.D. Advisors just need to understand the limitations of existing research processes and build a modern, more transparent process that leverages relevant technologies.

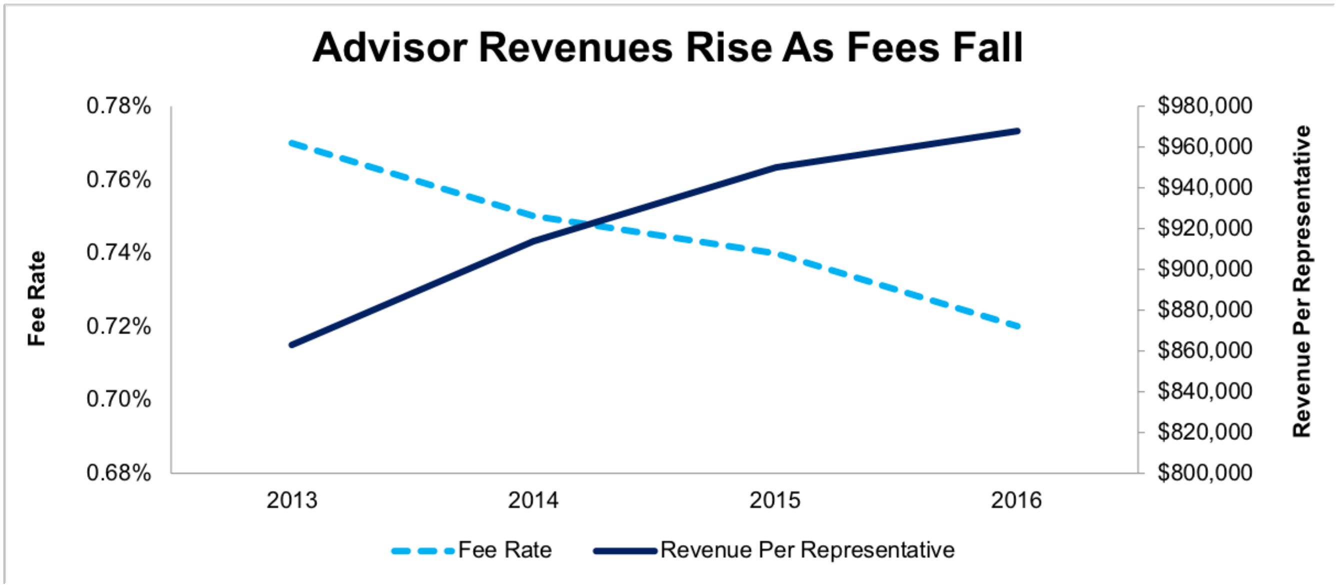
Advisors should think of AI technologies, like research automation, as partners that can automate large portions of manual processes. Automation frees advisors to serve a larger number of clients at a higher level.

AI Is Already Making Advisors More Money

Morgan Stanley (MS) is already successfully leveraging artificial intelligence to support its advisors. Like every other firm, Morgan Stanley has felt the impact of fee compression. Figure 4 shows that the rate on its fee-based client assets has fallen from 0.77% to 0.72% since 2013.

[5] Harvard Business School features the powerful impact of our research automation technology in the case New Constructs: Disrupting Fundamental Analysis with Robo-Analysts.

Figure 4: Morgan Stanley Fee Compression and Revenue Per Advisor



Sources: New Constructs, LLC and company filings

Despite these falling fees, revenue per representative has risen by 12%, driven by an increase in assets per representative. Since 2013, assets per representative are up from $116 million to $133 million, a 15% increase.

Morgan Stanley has used AI to automate administrative tasks and assist advisors with simple decisions. The company's "Next Best Action" system presents advisors with a range of suitable options for client portfolios, provides operational alerts, and even reminds advisors of notable events in their clients' lives.

Crucially, Morgan Stanley's AI empowers human advisors to do more, not to replace them. As Naureen Hassan, Morgan Stanley's chief digital officer, said in a speech last year:

"We are bringing in the latest developments in predictive analytics and machine learning — not to replace our advisers with some cyborg bot — but rather help them be faster and smarter in serving their clients' needs."

So far, Morgan Stanley and other wealth management firms have been mostly focused on the "faster" part of the equation. The next task is to use AI to make advisors "smarter" by leveraging its superior data processing power to provide better advice and fulfill the Fiduciary Duty of Care.

AI Enables Fulfillment of the Fiduciary Duty of Care

Fulfilling the fiduciary duty of care means showing clients and regulators that your investment advice is backed by research that is:

1. **Comprehensive:** Incorporate all relevant publicly available data (e.g. 10-Ks and 10-Qs), including the footnotes and MD&A.
2. **Objective:** Clients deserve unbiased research.
3. **Transparent:** Client should be able to see how the analysis was performed and the data behind it.
4. **Relevant:** There must be a tangible, quantifiable connection to stock, ETF or mutual-fund performance.

Most investors would be quite upset to learn that traditional research is far from meeting these qualifications.

Before AI, no human analyst or team of analysts could possibly perform comprehensive research[6] on their own for more than a handful of companies. The typical 10-K annual filing is over 200 pages of complex disclosures. Even if all you did was read 10-Ks all day every day, you'd never be able to keep up with the flood of documents filed with the SEC.

Historically, many advisors have trusted Wall Street to do comprehensive research. Unfortunately, the sell-side research model has some not-so-well-known flaws that make it unreliable in a fiduciary environment.

Advances in research automation technology now give advisors an alternative to traditional Wall Street research so they can faithfully fulfill the fiduciary duty of care for clients.

How AI Powers Research Automation

Our machine learning technology enables us to scale a "comprehensive" research process that big four accounting firm, Ernst & Young, proves is the best in the business. By automating large portions of the 10-K and 10-Q data collection process, we free up our human analysts for higher-value and more challenging work. In addition, our Robo-Analyst technology never tires and is not constrained by materiality thresholds. It collects everything all the time while also building highly complex financial models to QA and analyze the data. Meanwhile, our analysts focus on the most difficult and novel problems that the machine doesn't already know how to solve.

Our human analysts spend more time reviewing difficult disclosures in the management discussion and analysis (MD&A) like new deferred compensation plans for executives, changes in accounting practices, or unusual acquisitions and mergers. These are critical tasks that the machines do not know how to perform, yet.

For example, the Robo-Analyst recently flagged two items with a combined value of $322 million on Page 96 of PayPal's (PYPL) 2017 10-K. Analyst Pete Apockotos was able to quickly determine that these items—reversals of the loan loss allowance on a credit portfolio that was being sold—were non-operating income despite being included in the company's earnings and EPS. With the click of a button, Pete directed our Robo-Analyst technology to remove these items from our calculation of net operating profit after tax (NOPAT).

These non-recurring items accounted for 15% of PYPL's reported operating income. Any model that doesn't make an adjustment for them will significantly overstate the company's profitability.

We designed our AI to learn from every human analysts' insight. Once our human experts show the machine how to analyze data, the machine can do the work on its own. Since 2003, we've carefully seeded our AI and Robo-Analyst technology with millions of human analysts' insights from parsing over 120,000 10-K and 10-Q filings. Every day, our analysis of filings gets a little faster and a little smarter; so our analysts and our clients can spend more time focusing on more sophisticated activities than data gathering and model building.

This virtuous cycle allows us to continue to work together with machines rather than be threatened or replaced by them. The beauty of machines that learn is they enable us, humans, to do more.

How AI Can Help Advisors Grow and Keep Assets

AI becomes less frightening when you realize it's about expanding human capabilities, not replacing them. The robots are coming, and that's a good thing.

Transferring the slow, tedious work of reading 10-Ks to machines frees up analysts to spend time on higher-level strategic thinking and gives advisors more time to spend communicating with clients and better understanding their needs. Think of AI as an assistant that does the dirty work so you can be more productive.

[6] For example, institutional analysts, like Wells Fargo's Mike Mayo, leverage our research automation technology to enhance their work.

AI-driven research automation is a great tool for client retention, both as a value-added service and a way to educate clients about the logic and diligence behind recommendations. It's much easier to build buy-in on the investment process when you can be transparent about it.

Research automation technology also gives advisors more time to solicit new business, handle a larger client load, and serve clients at a lower asset level that has not previously been profitable. AI makes fiduciary-quality financial advice accessible to a larger number of people, which creates more potential clients.

It's true that AI will replace some human jobs, but as chess champion Garry Kasparov (who experienced firsthand being made obsolete when he lost to IBM supercomputer Deep Blue in 1997) puts it:

"Waxing nostalgic about jobs lost to technology is little better than complaining that antibiotics put too many gravediggers out of work. The transfer of labor from humans to our inventions is nothing less than the history of civilization."

Disclosure: David Trainer and Sam McBride receive no compensation to write about any specific stock, sector, style, or theme.

Visit us at www.newconstructs.com. Follow us on Twitter, Facebook, LinkedIn, and StockTwits for real-time alerts on all our research.

New Constructs® - Research to Fulfill the Fiduciary Duty of Care

Ratings & screeners on 3000 stocks, 450 ETFs and 7000 mutual funds help you make prudent investment decisions.

New Constructs leverages the latest in machine learning to analyze structured and unstructured financial data with unrivaled speed and accuracy. The firm's forensic accounting experts work alongside engineers to develop proprietary NLP libraries and financial models. Our investment ratings are based on the best fundamental data in the business for stocks, ETFs and mutual funds. Clients include many of the top hedge funds, mutual funds and wealth management firms. David Trainer, the firm's CEO, is regularly featured in the media as a thought leader on the fiduciary duty of care, earnings quality, valuation and investment strategy.

To fulfill the Duty of Care, research should be:

1. **Comprehensive** - All relevant publicly-available (e.g. 10-Ks and 10-Qs) information has been diligently reviewed, including footnotes and the management discussion & analysis (MD&A).
2. **Un-conflicted** - Clients deserve unbiased research.
3. **Transparent** - Advisors should be able to show how the analysis was performed and the data behind it.
4. **Relevant** - Empirical evidence must provide tangible, quantifiable correlation to stock, ETF or mutual fund performance.

Value Investing 2.0: Diligence Matters: Technology is Key to Value Investing With Scale

Accounting data is only the beginning of fundamental research. It must be translated into economic earnings to truly understand profitability and valuation. This translation requires deep analysis of footnotes and the MD&A, a process that our Robo-Analyst technology empowers us to perform for thousands of stocks, ETFs and mutual funds.

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 15 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.