Building The Last Hedge Fund — Introducing Numerai Signals

[Richard Craib](#)

[Follow](#)

Numerai

--

9

Listen

Share

Numerai is now the first hedge fund to source original stock market signals, built from any dataset, from anyone in the world. Numerai has allocated $50 million in cryptocurrency rewards for the most original signals.

Long before artificial intelligence destroys the world, narrow artificial intelligence will eat the stock market.

At Numerai, we've always felt it would be hard to build that artificial intelligence ourselves. Instead we've built a system for collecting stock market intelligence built by others.

Three years ago, I published [Numerai's Master Plan](#). Phase 1 of the plan was to monopolize intelligence

. And we have since built the largest stock market data science tournament in the world by providing free data and paying out millions to our community — over $40 million since launch.

In phase 1, Numerai's data scientists could use any machine learning algorithm to model the data we provided to create trading signals for our hedge fund. But for Numerai to contain all

stock market intelligence, we need to accept signals generated from any dataset

.

Today, we're announcing a new system that we've been working on for a long time. It's called [Numerai Signals

](https://signals.numer.ai) and it lets Numerai source and reward stock market signals built with any model, trained on any dataset.

It is the first step in phase 2 of the master plan: monopolize data

.

Where Is The Next Ken Griffin?

We have stock markets because it is valuable to society for stock prices to be right, to reflect all the information they can. If the prices are right, the best companies are getting the most capital to solve the problems in the world. It is really quite true that Tesla could not hire people, could not access capital, could not manufacture electric cars if its stock price was stuck at $1 and no one stepped in to trade it up to a fair price.

In 1987, the dream of people stepping up to correct market prices was alive. In his Harvard college dorm room, Ken Griffin raised $265,000 from his grandmother and his dentist so he could start trading stocks. He convinced Harvard to let him install a satellite dish on the roof of Cabot House to receive stock quotes. At 19 years old, Ken Griffin's insights were reaching and moving markets — in small ways at first and then in larger and larger ways as he proved himself.

In 2020, the dream is dead. Since 1987, many things have gotten better. Nearly every university student in the world has wireless internet that's one million times faster than Ken Griffin's 1987 satellite dish. And they have Robinhood on their smart phones that they can use to trade "for free". On the surface, it appears as though the playing field of the stock market has never been more level.

But in fact, it has never been more uneven. A bright young Ken Griffin in 2020 could get Robinhood but would soon realize that he's going up against the real Ken Griffin

who is still going strong as founder and CEO of Citadel: which manages $32 billion, has tons of data, and 1400 employees who are incentivized not to leave a scrap of money on the table for any of the young Ken Griffins. Not only that,

[but the real Ken Griffin is also buying up Robinhood's trade flow

](https://www.institutionalinvestor.com/article/b1m2p1cv68bx56/Twitter-Freaked-Out-Over-Robinhood-Selling-Its-Trade-Flow-But-the-App-and-Others-Have-Been-Doing-It-for-Years).

For a young Ken Griffin to have any hope at all he'd need to trade with a prime broker where he would need $10m of capital just to get started. To have any outside investors, he'd need a fund administrator, lawyers, auditors, etc. In 2020, the costs are just too prohibitive for starting a new hedge fund.

Pushed out of the real stock market with its high barriers to entry, little wonder the young Ken Griffins of 2020 would sooner start crypto funds than real hedge funds. At least in the crypto markets they could more credibly have an edge and all the costs would be lower. But that's not good. Having well-priced Dogecoins has a lot less

societal value than having well-priced Tesla stock. The magic of markets is in the consequence

of all the trading activity: in getting capital to the right companies. If all the young Ken Griffins of our generation become Dogecoin degenerates, the consequence is merely more Dogecoin memes.

But unfortunately, the high barrier to entry in starting a hedge fund is not the only problem. There's a big problem in how the stock market gives feedback as well.

A Feedback Problem Leads To Waste

One of the strangest things about working in quantitative finance is not being able to see the state of the art. If you're a car maker, you can test drive your competitors' cars or even buy one and dismantle it, and you can see exactly how your car compares. As a car maker, you can claim and prove definitive things like "our car has a top speed higher than the Porsche 911". Unfortunately, quantitative investment strategies are not like this at all. If you're a quant, you're in the dark about your competitors. You have no idea what they're doing. You have no idea if you're good.

You might develop a trading strategy that backtests well or seems to work in live trading but you have no idea whether what you have built is unique or better than the hedge fund across the street — and they definitely won't tell you.

Not knowing the state of the art is an infuriating and wasteful situation for the hedge fund industry. A startup hedge fund today might raise capital, hire employees and try to build a business around a trading signal that Citadel discovered in 2009 and Renaissance Technologies discovered in 1989 — and they'd never know it

. And if a signal has been discovered by others already, the extent to which the signal continues to work is only the extent to which Renaissance and Citadel are letting it work

.

A new hedge fund rediscovering a signal that is already known cannot possibly have a defensible or durable edge. In the long run, the market can't pay everyone for the same discovery, it can only pay participants who provide non-redudant

signals to the market. Unfortunately, for thousands of hedge funds, most of their signals are of the redundant kind: worthless because they're already discovered.

Trading against another fund when you have a strict subset of their signals is like choosing to play chess against Magnus Carlson whereyou start without a queen

. But many new hedge funds try to do just that. Convinced by often unscientific backtests, a continuous stream of new hedge funds without any differentiated signal end up food for Renaissance's monster because they had no way to know Renaissance had their signal already, and had them beat long before they even started trading.

So we have two major problems:

1. It's too expensive to start new hedge funds so good signals are stuck on the sidelines.

2. Even if you have started a hedge fund there's no way to know whether you have discovered a signal that has not already been discovered — no way to know whether you're providing a signal that's original and valuable to the market.

A bleak situation for the young Ken Griffins indeed. So bleak it almost makes you want to trade Dogecoin.

Not so fast.

The United States Signals Registry

Let's imagine a hypothetical new world without these problems. President Trump wants to take his mind off of COVID-19

and do something edgy before the election. He decrees that starting October 31st 2020 all US based hedge funds will have to submit their data to the United States Signals Registry ("Trump's USSR"). Every signal that every hedge fund uses for trading must be provided as a live feed to the USSR. All Renaissance's signals would need to be there, all Two Sigma's, all Citadel's, all George Soros'. Trump says it's going to be huge.

Trump wants all the signals so that the government can better understand the stock market and interfere with it more sensibly. But Trump's vision for the United States Signals Registry has hedge fund managers up in arms. They don't want to share their data with anybody, but Trump assures them that the data will be kept secret and that compliance is their only option.

In New York city, Jim Simons, the founder of Renaissance has an idea. He knows Trump is going to get his signals registry but Simons is not one to give up something for nothing in return. Jim Simons offers Trump an amendment to the way the USSR works: whenever anyone submits a signal to the USSR, they should get feedback on whether their signal is different from all the other signals in the system. Trump doesn't sweat the small stuff and immediately agrees with Simons' amendment.

The USSR is signed into law and established. Every hedge fund starts submitting their signals, and the USSR provides feedback on the extent to which a signal is different from all the other hedge funds' signals.

A few days after the launch of USSR, something powerful begins to happen in the stock market. Because market participants can now know before trading a signal whether or not another hedge fund has already discovered it, they stop wasting time on unoriginal signals. A golden age begins. Startup hedge funds clamor to upload new signals to the USSR to find out which of their signals are actually original. They focus on finding non-redudant signals — signals that no other hedge fund has found that could not possibly be integrated into stock market prices yet because they know that these original signals are the only durable edge they could have.

With market participants no longer in the dark about the other signals out there, a gold rush toward a new kind of market efficiency ensues. Quants begin to focus exclusively on new discoveries, integrating new types of data into their signal totally different from anything else in the USSR.

Some hedge funds are so demotivated by seeing how correlated their signal is to other signals in the USSR that they give up altogether, viscerally experiencing for the first time how little value they've been adding to the market their whole lives

.

Hedge fund investors begin to adopt a new investing policy. They refuse to invest in any new hedge funds based on backtests; they instead ask all new funds to first demonstrate that their signal is unique according to the USSR.

Data providers who are trying to sell data sets to hedge funds are told by the hedge funds to first submit their data through the USSR. Hedge funds refuse to even consider new datasets without first seeing the USSR's originality report on their signals.

Because of the credibility conferred to any signal with a positive result from the USSR, many hedge funds decide to shed 100% of their compliance, fund admin, legal, execution costs and instead sell their best, most original signals to other hedge funds. It seems peculiar to young people that we ever had 10,000 hedge funds in the first place with their duplicated datasets, unoriginal signals, paper pushing fund administrators, etc.

Because people trust the USSR, a new hyper-focussed industry of signal generators emerges. These signal generators only create signals and couldn't care less about their trading implementation. The young Ken Griffins become signal generators, and their insights now reach the real stock market and bypass the current barriers to entry. Interest in trading Dogecoin among young people collapses.

A few years later, Donald J. Trump and Jim Simons share the Nobel Memorial Prize in Economic Sciences for "for the establishment of the United States Signals Registry and its unreasonably effective contribution to stock market efficiency".

Although ashamed to admit it, even Hillary Clinton agrees the United States Signals Registry has led to better markets, far fewer wasted costs in the industry, and a far better world.

So what does any of this have to do with Numerai Signals?

Numerai Signals is the United States Signals Registry.

Numerai Signals

Numerai Signals is a new way to upload stock signals and figure out if your signal has predictive value after being neutralized by all the other signals Numerai already has including signals uploaded by others.

What doespredictive value after being neutralized

by all other signals

mean?

Consider the following example. Suppose Numerai has the P/E ratios of all stocks already (we do). This is a classic "value" signal used by many quants. Now suppose a new Numerai Signals user starts submitting his own custom value signal. This custom value signal is created with a neural network trained on P/E, Price to Book and Price to Sales ratios. He believes his custom value signal is much better and more predictive than simple P/E ratios so he uploads it to Numerai Signals.

Upon upload Numerai can see that his signal is very correlated with the P/E ratios we already have. In fact, even though he has put a lot of effort into his custom value signal, its correlation with standard P/E is 0.85. In lay speak, you might say Numerai already has 85% of what this new custom value signal is offering, so Numerai only cares about the performance of the 15% we don't already have.

Numerai Signals isolates the component of the signal orthogonal to P/E (and any other signals we have already) and rewards the creator of the signal based on that. By the predictive value after being neutralized by all other signals

, we mean the predictive value of this orthogonal component.

Numerai Signals is all about creating signals with predictive orthogonal components — the original part of the signals that we don't already have. The incentives are therefore around creating signals from unusual data sources or unusual modeling techniques. The feedback mechanism of Numerai Signals much more realistically captures what's valuable about a signal. Numerai Signals rewards people how the market should reward them: for the marginal predictive value of the non-redudant component of their signal.

By surfacing and rewarding only this non-redundant component in the signal, Numerai Signals can have the same impact on the world as the United States Signals Registry. It can allow signal generators to replace hedge funds. It can allow hedge funds to shed the burdensome costs of being money managers and become

signal generators instead.

Why use Numerai Signals?

In our hypothetical world when Trump proposed the United States Signals Registry, nobody wanted to share their signals. He only got the hedge funds submitting to the USSR because he said compliance was their only option. Numerai can't force anyone to use Numerai Signals, so why would anyone choose to use Numerai Signals instead of trading their signal themselves?

There are many reasons:

1. Most signals are not profitable enough by themselves to be worth trading.

(No one wants to invest in a hedge fund trading a signal that returns only 3% per year but this same signal would be very valuable to Numerai when combined with other signals).

1. Trading a signal properly in a hedge fund structure is prohibitively expensive but submitting to Numerai Signals is free.

2. It's unwise to trade against other funds that may already have your signal and many more as well.

3. When you trade on the market, you can't tell if you're trading an original signal. On Numerai Signals, you can see exactly how original you are.

4. Nothing submitted to Numerai Signals is exclusive. You can trade a signal yourself and submit it to Numerai Signals as well.

5. To earn money on Numerai Signals, you don't need to trade anything instead you simply submit your signal and stake it with the NMR cryptocurrency just like on Numerai. The minimum stake is only 0.01 NMR ~30 US cents.

6. On Numerai Signals you can evaluate the originality of your signal on historical data as well. These historical diagnostics give fast feedback on the originality of your signal before you stake it.

Examples of Signals Users

An intern at PayPal strings together some transaction data and comes up with what he believes is a strong signal that works to predict the returns of tech stocks. He approaches PayPal's CEO and says "I think we could start a hedge fund with this". PayPal's CEO says, "our business is not trading; trading is a complicated, expensive business". That evening the intern discovers Numerai Signals and sends an email to the CEO: "we can make 200% per year on my signal built from our data without doing any trading at all". PayPal's CEO is intrigued and agrees to let the intern try it out with a small stake. Because the intern's signal is so good and original, PayPal decides to build a whole new business unit around submitting signals to Numerai. They earn large payouts just by submitting a small csv file to Numerai's API each week.

A New Jersey based global equity hedge fund has been banging its head against the wall trying to get their quant strategy to work. Over the last few years, performance has been weak. Last year, they made 2% but spent 8% on trading costs (short borrow costs, market impact costs, etc). If they didn't have to pay the trading costs their signal would earn 10% per year. They consider selling their signal and technology to a hedge fund but every hedge fund they approach says "if your signal is so good, why don't you trade it yourself". A quant at the firm gets permission from his boss to upload the signal to Numerai Signals. He immediately discovers their signal is highly orthogonal to everything else on Numerai. Their hard work really has meant something. He tells his boss that their signal is original but they may never be able to cover their costs trading it in the real stock market. They decide their best move is to submit their signal to Numerai Signals where they pay no trading costs. For a while, they continue trading the signal as well but eventually figure out they can make a lot more money a lot easier just by submitting to Numerai Signals. Previously, their only option was to close their fund and deprive the market of the valuable signal they'd worked on for so long but now their signal gets full expression in the market through Numerai Signals.

Jimmy has been a Quantopian user for several years. He has taught himself to be a quant on Quantopian. He has run 27,000 backtests, and he has even built a few signals that ran in the Quantopian hedge fund before it shut down. Despite his skills and performance on Quantopian, Jimmy has no way to get full value for his signals. He asks Quantopian permission to allow him to automatically submit his signals to Numerai Signals. Quantopian likes the idea and agrees to let their entire community submit their signals to Numerai Signals automatically through the API. Quantopian strikes a deal with Numerai to share in the upside.

There is little doubt in Jason Rosenfeld's mind that modern language models like OpenAI's GPT-3 are the future. He's convinced that AIs can basically comprehend text at this point. He has an idea to use language models to create new, original stock market signals. Jason Rosenfeld has been working on NLP for years and used to work at $38 billion hedge fund, Millennium but decided to start his own company, CrowdCent. Although, Jason has been modeling Numerai's data for a while, there has been no way for him to use Natural Language Processing techniques like those used by GPT-3 on Numerai's data (it's numerical data, not set up for language models). But CrowdCent has gigabytes of text data referencing different stocks which Jason has already built signals on. Jason knows his signals are institutional-grade, and capable of powering trading strategies at the likes of Millenium but Jason knows he can make more money submitting his signal on Numerai Signals. This is a real example.

Jason is currently in first place on the Numerai Signals leaderboard and he is currently building out a team to stay number one.

If you're like any of these examples,

[sign up to chat with us

](https://community.numer.ai/channel/signals)or

[read our technical documentation

](https://docs.numer.ai/numerai-signals/signals-overview).

All The Signals

Numerai Signals is a way for Numerai to collect entirely new complementary signals based on any new data sources. It is an extension of Numerai's API. You can sign into Numerai Signals with your Numerai account, stake with your existing NMR wallet balance, and use the same API keys to submit. Just like Numerai, it is built on the Erasure protocol and uses the Numeraire (NMR) token for staking and payouts. Even the universe of ~5000 stocks is exactly the same between Numerai and Numerai Signals.

The reason Numerai Signals is so tightly coupled with Numerai is because its purpose is to gather signals which enhance Numerai's core signal — our meta model which combines all the models staked on Numerai.

Any data scientist without data can create signals based on Numerai's data and continue to improve the meta model. But now any data provider, quant, or even other hedge funds with their own data can use Numerai Signals to see how original their signals are and stake them if they believe in them.

Over time the goal is that Numerai Signals becomes the best place to monetize all stock market signals. With there finally being a way to know whether your signal is original, you can now focus your efforts on that piece alone. Instead of having 10,000 hedge funds each duplicating each others strategies in the dark, we can finally have consolidation in the industry, with one peaceful, open hedge fund connected to every signal.

Numerai Signals is a major next step for Numerai's master plan so we have decided to allocate $50 million of our cryptocurrency treasury for rewarding new signals. Compared to Two Sigma's financial modeling competitions on Kaggle ($200k) and all Quantopian's contests ($180k), Numerai Signals rewards are 100 times larger

making us the most important quant community on the planet. Join us.