Hedge Fund Law Report
An Acuris company

Lowenstein Sandler

January 9, 2020

TECHNOLOGY

Key Compliance Considerations for Fund Managers Using Alternative Data

By William V. de Cordova, *Hedge Fund Law Report*

Over 80 percent of hedge funds are using alternative data, including biometric data, geolocation data and web scraping, according to a recent survey conducted by Lowenstein Sandler. Although use of alternative data is expected to increase, stronger privacy regulations – such as the recently enacted California Consumer Privacy Act of 2018 (CCPA) – will affect fund managers' ability to source and use that data.

To explore how fund managers can become comfortable using alternative data, the Hedge Fund Law Report recently spoke to Peter D. Greene, partner at Lowenstein Sandler and author of the survey. This article sets forth Greene's insights on the key compliance issues raised by using alternative data, including insider trading and privacy concerns; new and prospective regulatory issues in the U.S. and abroad; best practices for mitigating risk and managing third-party data providers; and ways newer forms of alternative data are affecting fund managers.

See our three-part series "A Fund Manager's Roadmap to Big Data": Its Acquisition and Proper Use (Jan. 11, 2018); MNPI, Web Scraping and Data Quality (Jan. 18, 2018); and Privacy Concerns, Third Parties and Drones (Jan. 25, 2018).

To further explore these issues, on Wednesday, January 15, 2020, at 11:00 a.m. EST, the Hedge Fund Law Report will host a complimentary webinar, entitled "Best Practices for Private Fund Managers' Use of Alternative Data." Moderated by William V. de Cordova, Editor-in-Chief of the Hedge Fund Law Report, the panel will feature Adam Reback, director at Optima Partners; Stacey M. Brandenburg, shareholder at ZwillGen; and Jeffrey Neuburger, partner at Proskauer. The discussion will address issues including the pros and cons of generating or purchasing datasets; managing third-party data providers; complying with data privacy laws and cybersecurity guidance; and avoiding insider trading and other risks. To register for the webinar, click here.

HFLR: What are the key compliance issues that arise when fund managers use alternative data?

Greene: There are two that are most important – one far more than the other in a sense.

First is making sure that, from a securities-law perspective, you're not walking into an insider trading problem. The second is privacy, which is now a complex array of state, federal and international laws that is very difficult – if not impossible – to navigate.

Thus, while we can drill down on these a bit more, the two concerns fund managers should worry about most when buying and consuming alternative data are insider trading and privacy.

HFLR: With respect to insider trading, what are the primary regulations and other considerations that private fund managers need to keep in mind when using alternative data?

Greene: For insider trading, the primary regulations are the U.S. insider trading laws – Section 10(b) and Rule 10b5-1 under the Securities Exchange Act of 1934 (Exchange Act).

When thinking about insider trading in the U.S., you think about three elements to the crime:

1. Is the information material?
2. Is the information nonpublic?
3. How did you obtain it? Did you either misappropriate it or receive it in breach of a duty?

[See "HFLR Panel Identifies Best Practices for Avoiding Insider Trading Liability in the Aftermath of *Martoma*" (Jan. 18, 2018).]

With respect to every type of alternative data except for web scraping – by that, I mean credit card transaction data, social media sentiment data, app usage, geolocation, satellite imagery, all these different kinds – there is a pretty good argument that the information is material, because the fund manager is spending money (often a significant amount) to obtain that data. So, that satisfies the first element. We know that the data is not public; otherwise, why would the manager be paying for it if it were otherwise available in the public domain? That satisfies the second element.

Thus, we know that the government likely can meet the first two elements with respect to insider trading in the U.S. for nearly every type of alternative data, except web-scraped data. How, then, can a manager become comfortable using that data?

The way for a manager to become comfortable is to be very careful about data provenance. In other words, we want to make sure that there is permission at every link in the chain from the original creator/owner of the data – if it's credit card data, from the credit card user's use of the credit card, or if it's geolocation data, from the use of a person's phone and the way it tracks where the person goes – to the hedge fund manager buying the data.

We are able to do that in most, although not all, instances by using a good due diligence questionnaire (DDQ); conducting careful diligence with the vendor; and negotiating a comprehensive agreement with robust representations and warranties. Once we're able to do that, then we're comfortable from a U.S. perspective that we can buy the data.

The one exception I mentioned was web-scraped/web-crawled data, because with that data, we know that the website has said in nearly every instance that scraping or crawling is prohibited. So, how can a fund manager become comfortable buying web-scraped data – or web scraping on its own? Under the recent decision in *hiQ Labs, Inc. v. Linkedin Corp*, the U.S. Court of Appeals for the Ninth Circuit held that, even if someone is scraping data from a website that says that scraping is prohibited, as long as the data is "public" (*i.e.*, you don't have to put in a password to access it), the scraping is permissible (although the case was not an insider trading case). Thus, with respect to web-scraped data, there isn't going to be an insider trading problem because that data is considered public, and in order for insider trading to exist in the U.S., the information in question has to be nonpublic.

HFLR: That covers the U.S. What about in other jurisdictions, such as the U.K.?

Greene: Although I do not practice U.K. law, I know from my work with Leonard Ng of Sidley Austin that the law is different in the U.K. Specifically, in contrast to the U.S., in the U.K. (and all of the E.U. for that matter), there are only two key elements to the crime of insider trading using "inside information":

1. materiality – which they call "price sensitivity"; and
2. nonpublic versus public.

There is no element in the U.K. related to misappropriation or breach of duty. So, what does that mean? It means that the issue comes down to the question of whether the data is publicly available.

In the U.K., the definition of "public" is a bit different than in the U.S. In the U.S., public is generally viewed as widespread dissemination. In the U.K., however, it's largely widespread availability. As a result, if you can buy the data set, that likely is enough for the information to be deemed public. Thus, most of these data sets that are allowed in the U.S. are also allowed in the U.K.

In essence, there are well-established laws in the U.S. and other countries regarding insider trading, and what we are doing on the alternative data front is simply applying those laws to a new set of facts. Just like we applied insider trading laws to expert networks more than 10 years ago and then applied them to political intelligence firms more recently, we are now applying those same laws – that don't expressly discuss alternative data – to a new set of facts. And, as we do so, we must examine what the alternative data looks like, what the different data sets are, etc.

[See also our three-part series on what fund managers need to know about corporate access: "The Risks and Rewards of Speaking Directly With Issuer Management" (Nov. 15, 2018); "Six Front-End Controls to Manage the Risk of Inadvertently Receiving MNPI" (Nov. 29, 2018); and "Implementing Testing and Preparing for SEC Scrutiny" (Dec. 13, 2018).]

HFLR: What are the primary regulations and other considerations that private fund managers need to keep in mind from a privacy standpoint when using alternative data?

Greene: Privacy is, one might say, a morass of many different laws. You have the Gramm-Leach-Bliley Act (GLBA); you have the new CCPA; you have the E.U.'s General Data Protection Regulation (GDPR).

[See our two-part series "Engaging With the California Consumer Privacy Act": How Hedge Fund Managers Can Evaluate Whether They Are Subject to the New Law (Sep. 26, 2019); and How Hedge Fund Managers Can Prepare for Compliance With the Act (Oct. 3, 2019). See also our two-part series on the GDPR: "Impact" (Feb. 21, 2019); and "Compliance" (Feb. 28, 2019).]

It is becoming very complicated for hedge fund managers to understand how all these different laws fit together. Hopefully, there will eventually be a federal law in the U.S., which I believe would be welcomed by many. For now, however, managers need to piece through all of the various state laws.

What that means for a manager in the alternative data context is that it needs to make sure, when buying a data set, that it conducts appropriate diligence and that it obtains representations regarding personally identifiable information (PII) in the U.S. and personal data in the E.U. The main questions a fund manager should be looking to answer are:

- Is it receiving any PII or personal data?
- Is it receiving anything that might enable the manager to reverse engineer or back into the identity of a particular person or data point related to a person?

The answers to both of these questions should be "no."

Once a manager is comfortable that it will not be receiving any PII or the ability to reverse engineer any PII, and actually buys the data, the manager needs to then inspect that data and make sure that it did not, in fact, receive any PII. If the manager does inadvertently receive PII, despite the fact that the third-party data provider has contracted with it and covenanted that it would not provide the manager with that type of information, the question then is what should the manager do about it?

A manager in that situation has two choices:

1. delete the PII and maintain a deletion log. This is tricky, however, because the Investment Advisers Act of 1940 (Advisers Act) requires registered investment advisers to keep certain books and records of the investment adviser; or
2. sandbox the PII, putting it onto a special, segregated server to which only legal, compliance and information technology personnel have access.

HFLR: As you have pointed out, privacy requires compliance with various laws, including the GLBA, the CCPA and the GDPR, which was recently amended. Do you expect to see any big cases or enforcement actions about alternative data in the near future?

Greene: There hasn't been an insider trading case around big data other than one Capital One case many years ago. That wasn't really a hedge fund case, however; that case just involved two rogue Capital One employees.

Although there hasn't yet been an insider trading big data case, at some point there will be, especially with how much exposure privacy generally is given in the mainstream press. Seemingly on a weekly or monthly basis, Apple, Facebook and Google are slugging it out to show who cares more about consumers' privacy. Privacy has become such a hot-button, mainstream issue that I will not be surprised if we soon wind up seeing a case involving the exploitation of data by a fund. Like all the other first movers in this area, however, I believe that it will involve an egregious set of facts.

[See "How the GDPR Will Affect Private Funds' Use of Alternative Data" (Jun. 14, 2018).]

HFLR: How can fund managers try to avoid becoming subject to such a case?

Greene: What I see all my clients do is take very seriously the compliance process. They put in place a DDQ. They diligence data vendors before they start to negotiate an agreement; they or their outside lawyers have a call with the vendor to ask questions around the vendor's compliance policies and procedures with respect to the collection of the data they will be purchasing. Then, once they are comfortable from a diligence perspective, they start to negotiate an agreement that contains robust representations and warranties around data provenance and the absence of PII.

[See "Best Practices for Due Diligence by Hedge Fund Managers on Research Providers" (Mar. 14, 2013); and "Hedge Fund-Specific Issues in Portfolio Management Software Agreements and Other Vendor Agreements" (Aug. 4, 2011).]

HFLR: Other than conducting a call with a vendor to ask about compliance procedures with respect to data provenance, what other best practices should managers follow to vet providers of alternative data?

Greene: Some of this will be repetitive, but the first is that every manager must have a DDQ. We work with each of our clients to design a DDQ that fits its business.

The next step is generally to have a phone call, either conducted by in-house or external counsel, with compliance at the vendor to ask questions and drill down on certain things in the DDQ. How do you get your data? From where do you get it? How are you comfortable that it's "clean" – meaning that the vendor is allowed to have it and, just as importantly, that it is allowed to sell it?

In some instances, we have had success convincing vendors to turn over to us their contracts with their data suppliers. The economics of the contract are redacted, of course, but the contract still shows that, in fact, the vendor is allowed to buy the data it is buying and allowed to sell it to us for the purpose for which we want it – which is to make investment decisions.

Once we complete the due diligence process, we then negotiate a solid contract. As I mentioned, that contract must contain robust representations with respect to data provenance and privacy.

[See "Fund Managers Must Supervise Third-Party Service Providers or Risk Regulatory Action" (Nov. 16, 2017).]

HFLR: What about after the contract is signed? What must a fund manager do?

Greene: When the data comes in, a manager must make sure that the data is scrubbed so there is no PII and that the data that the manager thinks it is buying is, in fact, the data it is receiving. The manager should then conduct periodic diligence tests and checks on data sets. It's not realistic for compliance to look at every data set that comes in, but it would be nice for compliance to periodically scrub and review the data.

HFLR: Do you think that the average compliance officer – who may not have a strong technical background – will have difficulty staying on top of this?

Greene: That is a fair question, and there certainly is a limit to what legal and compliance professionals can do with respect to understanding exactly how data was captured and sourced. I personally don't have a technical background.

That being said, if you ask the right questions, the absence of a technical background should not be a barrier.

HFLR: What would you say some of those questions are?

Greene: The first question would be, "Walk me through, in layman's terms, where you get the data from; how that person or party gets the data; and whether any electronic means are used to source, acquire or review the data."

A follow-up question would be, "If electronic means are involved, do any of those means use any masking or obfuscation as to your identity when sourcing, acquiring or reviewing the data?"

HFLR: In a way, could a lack of a technical background be an asset to a certain degree, as it would force a legal or compliance person to ask very basic questions?

Greene: The above are essentially very basic questions that should not require a lot of big words to answer. If big words are used in the answer, I (or the relevant legal or compliance person) may not understand what the other person is talking about, so it is much more helpful for the vendor to explain everything in lay terms.

HFLR: Fund managers have been using consumer data such as credit card transactions for years. How have newer forms of alternative data, such as biometric and geolocation data, changed the scene?

Greene: Biometric and geolocation data are worrisome, because they are, arguably in many instances, personal information. Thus, managers have to be even more careful with biometric and geolocation data to ensure that they cannot reverse engineer the identity of the person whose biometric or geolocation data they are using. As a result, while fund managers are buying and using biometric and geolocation data, if anything, they will be even more careful with their diligence from a data-provenance perspective.

As we see from our firm's recent survey and report, however, managers are very interested in these types of data, and they expect to look at and buy more of those types of data sets going forward.

The two most interesting takeaways from the report to me, as someone who does a good deal of work in this space, are:

1. the interest in biometric and geolocation data; and
2. the fact that approximately 98 percent of managers that are buying data are using it in combination with traditional fundamental analysis.

The stories we read years ago that suggested that alternative data would replace the fundamental investment process were greatly exaggerated. Managers are not simply buying data sets and making decisions based on those data sets. Rather, they are buying data sets and combining them with their fundamental analysis to make investment decisions. We are seeing a lot more of this so-called "quantamental" – or really "data-mental" – analysis.

[See "Best Practices for Private Fund Advisers to Manage the Risks of Big Data and Web Scraping" (Jun. 15, 2017).]

HFLR: What could explain the increasing interest in biometric and geolocation data?

Greene: There is so much more geolocation data. Think about every app you have on your phone and how your movements are tracked. It's just proliferating. Therefore, it's logical to see funds buying more of that data because more of those data sets exist.

As for biometric technology, I think we're in the early innings of the biometric technology game, but the same explanation applies to it as to geolocation data.

HFLR: Has the rise of alternative data actually improved the investment process?

Greene: I don't think we know the answer to that yet. It's too early to know that, and it's also hard to gauge that. Once we have significant historical information with the ability to compare the performance of purely fundamental shops against that of data-mental shops, we will have a better sense of how much alternative data has influenced the process.