



K&L GATES

2017 INVESTMENT MANAGEMENT CONFERENCE  
NEW YORK

# Big Data: Risks and Rewards for Investment Management

Derek N. Steingarten, New York

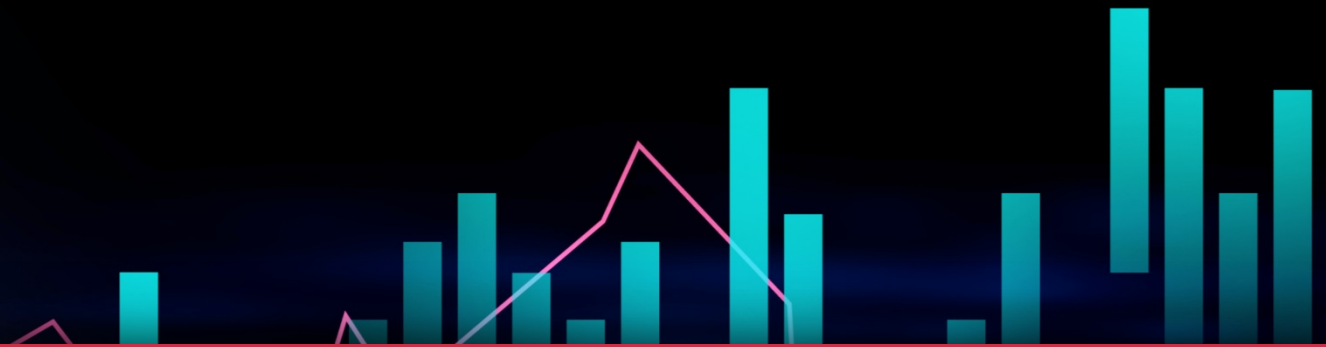
Julia B. Jacobson, Boston

Barbara Bridges, VP of Legal & Compliance, Demystify Data

October 18, 2017

## OVERVIEW OF PRESENTATION

- What is “Big Data?”
  - How are funds using it to enhance analytics and trading?
- Legal and Compliance Issues in Sourcing Quality Data
- Vendor Due Diligence
- Liability and Contracting
- Insider Trading
- Privacy and Cybersecurity Compliance



# 'Big Data' Basics



## WHAT IS BIG DATA?

Federal Trade Commission (Jan 2016): “nearly ubiquitous collection of ... data from a variety of sources, the plummeting cost of data storage, and powerful new capabilities to analyze data to draw connections and make inferences and predictions.”

- **Volume** – quantity of data that can be gathered and analyzed
- **Velocity** – speed at which data can be gathered and analyzed
- **Variety** – combinations of very different, once unlinked data to infer and predict
- **Veracity** – accuracy of data and analysis



## HOW DOES BIG DATA DIFFER FROM 'TRADITIONAL' DATA?

- Data type
- Data source
- Data privacy and data security risks



# HOW DOES BIG DATA DIFFER FROM 'TRADITIONAL' DATA?

- Data Type - **Personal Data**
  - Data Breach Notification Laws
  - GALBA: Nonpublic Personal Information
  - FTC: “not yet linked to a particular consumer, computer, or device but that may reasonably become so”
  - EU: Personal data is data “relating to an identified or identifiable natural person; an identifiable natural person is one who can be identified, *directly or indirectly*, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person”



## HOW DOES BIG DATA DIFFER FROM 'TRADITIONAL' DATA?

- Data Type (cont.)
  - **Anonymous vs. Anonymized vs. DE identified Data**
    - All terms of art
  - **“Pseudonym zed Data” (EU term)**
    - Neither anonymous nor directly identifying
    - Separation of data from direct identifiers so that linkage to an individual is not possible without additional information that is separately held
    - Reduce the risks associated with data processing while also maintaining the data's utility

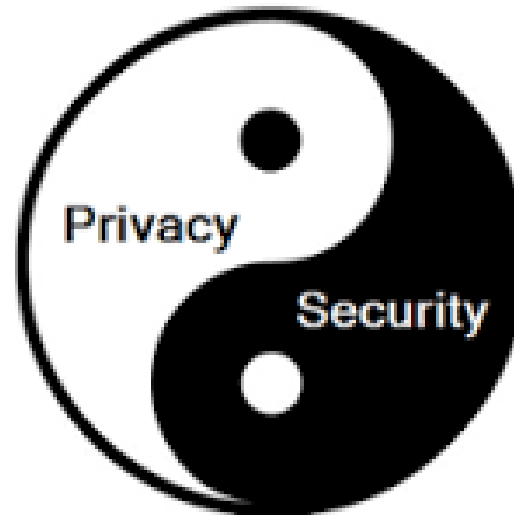




# HOW DOES BIG DATA DIFFER FROM 'TRADITIONAL' DATA?

Information privacy is about **Individual Rights**

Notice  
Choice  
Accountability  
Purpose/Use Limitations  
Access (correct, amend, delete)  
Recourse & Liability



Information security is about **Protecting Information and Information Systems**

Authentication  
Access & Use Controls  
Encryption  
Storage  
Disposal  
Incident response



## “PERSONAL” VS. “CONFIDENTIAL”

- ❖ **Confidential information** may include personal information.
- ❖ **Confidentiality** obligations are not sufficient to address **privacy** and **security** obligations.

**YES:** **security** without **privacy** or **confidentiality**

**NO:** **privacy** or **confidentiality** without **security**



# HOW DOES BIG DATA DIFFER FROM 'TRADITIONAL' DATA?

## DATA QUALITY & ACCURACY



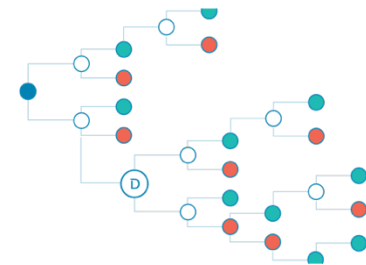
- Are unstructured and structured data matched and compared against each other?
- Is it the most relevant type of data?

## DATA SECURITY



- What information is sent or received?
- How is it protected against potential cyber threats?

## DATA AUTHENTICITY



- Is the data vendor “legitimate”?
- Does the vendor have good data hygiene?
- Can the vendor demonstrate that the data was lawfully sourced?



# How and Why of Big Data

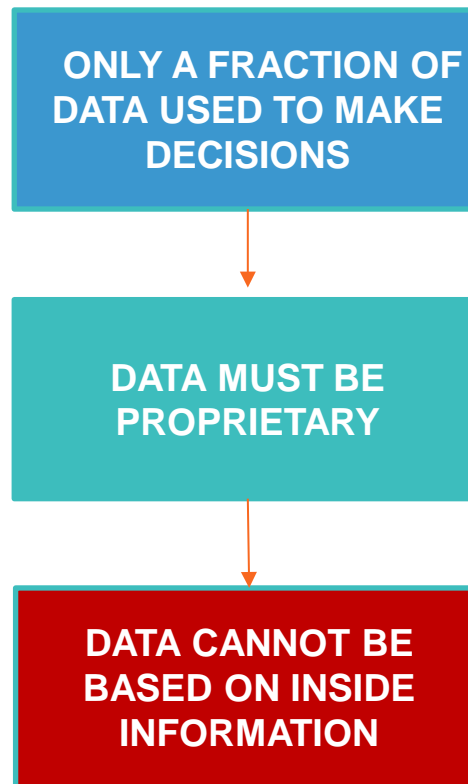


## HOW CAN HEDGE FUNDS AND OTHER INVESTMENT FIRMS USE BIG DATA?

- Quantitative strategies
- Fundamental analysis
- Trading
- High level trends analysis and investment decisions making, more depersonalized
- Consumer data, more personalized data (e.g., credit information) – privacy and data protection risk
- Machine learning
- Predictive analytics



# LEVERAGE BIG DATA – BETTER DECISION-MAKING



## BENEFITS TO IMPROVED DATA AGILITY & MANAGEMENT



### Focus on Opportunities from Dynamic Modeling

- The key to a competitive edge is not in the proprietary nature of the data but the dynamic modeling (e.g., track changes and set triggers)



### Enhance Data Agility and Discovery

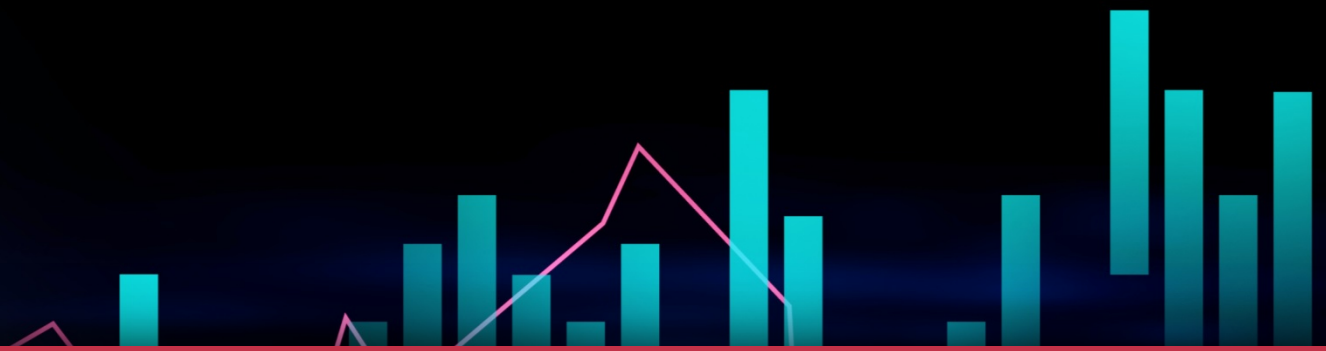
- Improving data agility and speed generates improved discovery - new signals and insights



### Strengthen Coordination and Data Management

- A 360-view of data usage leads to better coordination between data management functions and enhanced insight sharing across the business





# Sourcing Big Data





# BUILD IT IN HOUSE OR BUY IT FROM VENDORS?

## Dealing with Third Parties

- License vs. Acquisition - when does use vs. ownership matter?
  - Derivative works?
- Risk Allocation
  - Privacy
  - Cybersecurity (SEC, Equifax)
  - Data quality / service quality
  - Indemnification
  - Liability allocation (cap?) / representations



## BUILD IT IN HOUSE OR BUY IT FROM VENDORS? (CONT'D)

### Confidentiality

- Notice process?
- Protecting proprietary investment strategies

**Scraping and crawling** - systematic extraction of data accessible on a web site.

- Is it legal?
- Would you want it on your website?
- Insider trading concerns (MOPE)?





# Big Data Maintenance



# HOW TO ON-BOARD AND MAINTAIN BIG DATA

## Data governance issue

- Cloud concerns
- Tracking/monitoring access: who / why / how
- Cross-border issues
- Machine learning

## Trading compliance: Insider trading and MOPE concerns

- Risk that the vendor will access manager's system to download information and strategies and monetize
- How to mitigate risk





Questions?



K&L GATES