

No AI Without Big Data

# Key Points **History of Big Data**



**The Beginning of the Information Overload**

8 years to tabulate  
**Tabulate US Census Data**

1880

1881

1900

1920

1930

1940

1960

1964

1993

1994

1995

1997

1999

2000

**Economic Boom**



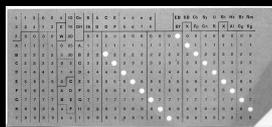
The rise of **Centralized Computing Systems**



World Wide Web growth amplifies big data problem



Taming the "big data"  
**Invention of Punch Cards**  
The Hollerith Tabulating Machine



**The Impact on Libraries**



**Internet of Things** coined



The term "big data" was first used by NASA researchers  
**The Problem of Big Data**

## Gartner's 3Vs

Volume, Velocity, & Variety



## The Launch of AWS

2001

2002

2003

2004

2007

2012

2014

2015

2016

2017

2018

2019?

## The Facebook

appears with the rise of social networks



Facebook reaches 1 billion users



## Big Data

influences the outcome of US election?

Serverless Computing on the rise

## Birth of the "Alicorns"

Apple Inc. & Amazon touch \$1 trillion mkt cap.



RavenPack is born. First alternative data for quant firms



Data science, fake news can be used to **manipulate public perception**

Data Privacy  
GDPR, the Cambridge Analytica data scandal, global regulatory scrutiny...



# Big Data **in Asset Management**



**80%**

of available data to hedge funds is  
unstructured and most of it is text

# Generating Alpha with **Unstructured Data**



## **Text**

News, filings, social media, research reports, emails, chats, etc.



## **NLP Engine**

Text parsing, named entity recognition, tagging, sentiment, novelty & temporal analytics



## **Analytical Tools**

Statistical analysis, machine learning, & predictive modelling



## **Alpha Generation**

Alpha signals, input factors, & overlays

# NLP Engine



## Entity Detection

Named entity recognition of people, places, and things



## Relevance

Identify the importance of entities mentioned



## Topic Tagging

Determine what documents are about



## Parsing

Understand the meaning of words and sentences



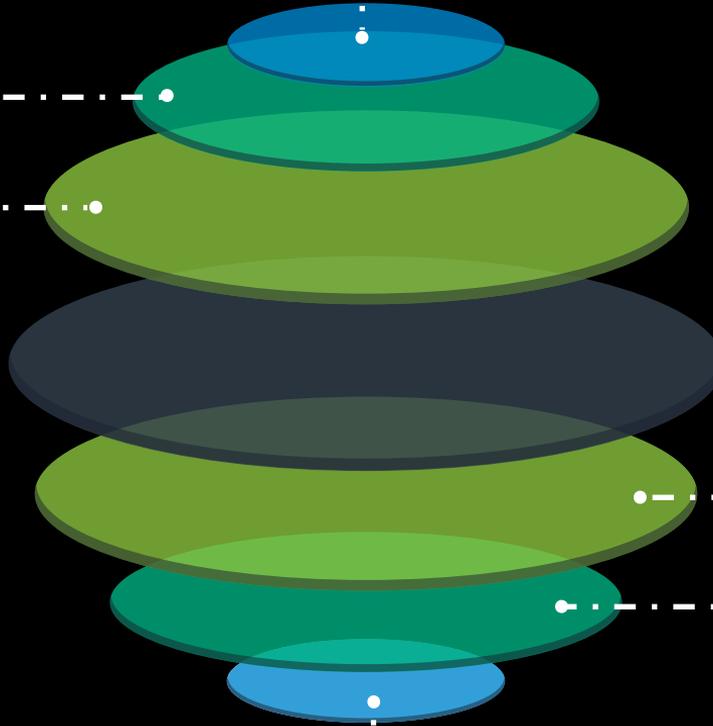
## Novelty & Temporal

Similarity analysis and temporal data

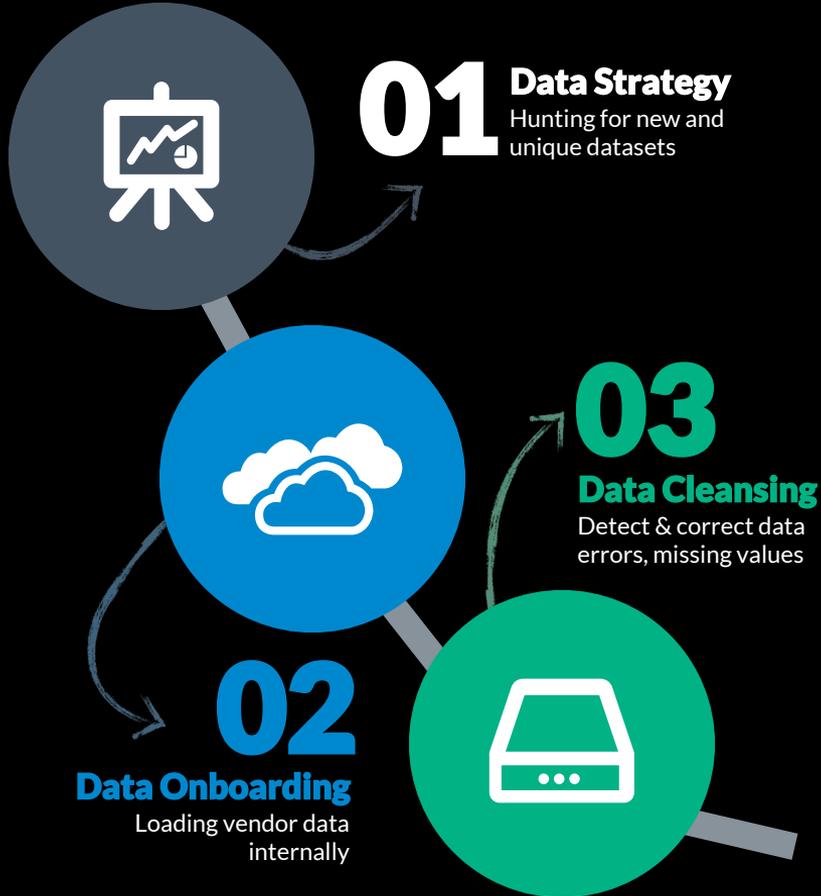


## Sentiment Analysis

Measure positive, negative, or neutral sentiment



# Big Data @ AI Funds



All about the data...

# Big Data @ AI Funds



## Build or Buy?

Evaluate true cost of sourcing multiple data feeds

## Raw or Machine Readable?

Unstructured data cannot be handled by traditional data tools

## DIY or Hire?

Hard painstaking tasks that take up valuable time and resources

All about the data...

# Big Data @ AI Funds



# Big Data @ AI Funds



# Big Data @ AI Funds



# Big Data @ AI Funds



No AI without Big Data