



Accelerating Innovation in the Era of Big Data

Dr. Lucas Brown
RedOwl Analytics



My background

- PhD from Oxford University
- Senior Data Scientist with President Obama's re-election campaign
- Senior Data Scientist with **RedOwl Analytics**, building a software platform for big data analytics



What is **Big Data**?

- **Large** volumes of data
 - Larger than `traditional' database methods can handle, so a few terabytes+
- **Live** ingest
- **Analytics** -- simple to complex



Opportunities from **Data**

- Extraordinary growth in digital data
 - All aspects of the firm are digitized
 - Purchases, inventories, email, phone logs, energy
 - Often digitized for other reasons, but now provide additional value opportunities
 - **40% growth** in global data generated per year
 - Retailers using big data can increase operating **margins by more than 60%**



Opportunities from **Data**

- Solid products that make distributed computing accessible
 - Amazon's **AWS**
 - Apache **Hadoop** and MapReduce
 - Increased **security in cloud deployments**



Challenges from **Data**

- Companies are buried in data without training to understand it
 - **Only 5% growth** in IT spending per year
 - US faces a **shortage of 1.5 million managers** and analysts to make decisions based on analytical findings
 - 0.5% of the US population
- Data is spread out over **multiple unconnected systems**



Challenges from **Data**

- Software products **don't always help**
 - **Too many alerts**
 - Increased liability





Solutions from Data

INTEGRATE DATA

Communications

- Email
- Chat
- IM
- Bloomberg
- Phone Calls/Logs

Transactions

- Trades
- Account Movements
- General Transactions

Physical Meta Data

- Door Swipes
- Building entrances

Computer Activity

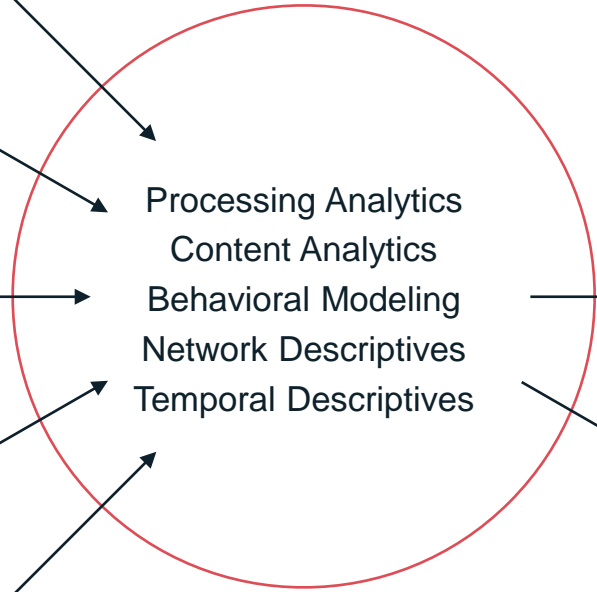
- Web Activity
- Printer Actions
- File Activity

Enrichment Data

- HR system data

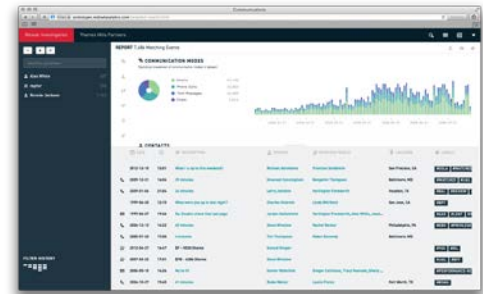


CALCULATE ANALYTICS

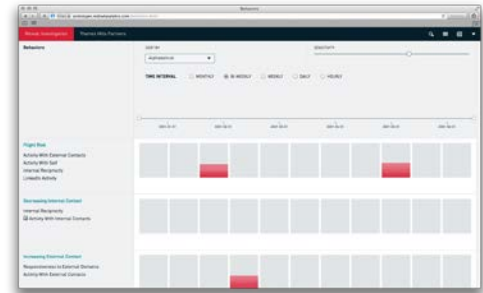


VISUALIZE

Search & Investigate



Daily Monitor & Alert





Solutions from **Data**

- Presenters today will talk about practical ways to extract **dollars from data**
 - From **retail customers**
 - From **machine data** and network activity logs
 - From **energy data**