Spark'ing an Anti-Money Laundering Revolution

Katie Levans | Koert Kuipers | Tresata

Overview

The Thin Line Between Good & Evil
Why Money Laundering is Hard to Stop

Lots of Rules, Not Enough Tools
Why AML is Top of Mind (But Hard to Do Well)

Spark'ing a Revolution
Tresata + Databricks Solution

TOO BIG

$2 TRILLION PROBLEM

tresata


TOO BROAD
FOOD OR GUNS
tresata


TOO DUMB
98% false positives
tresata

tresata


AML IS A NETWORK PROBLEM
tresata

Money Laundering Prevention

Customer Due Diligence (CDD)
Requirements to Know Your Customer by naming each account holder & proving identity

Reporting
Requirements for reporting of information to regulators

Regulation & Supervision
Countless rules and governing bodies to keep up with

Sanctions
Financial punishment for failing to uphold the first three measures

AML Fines

$10 million Ocean Bank (2011)

$1.2 billion HSBC (2012)

$2 billion JP Morgan Chase (2014)

Source: http://www.bankersonline.com/security/bsapenaltylist.html#citibankna

Problems with Current AML Tech

Poor entity resolution and identification

Tedious manual investigation

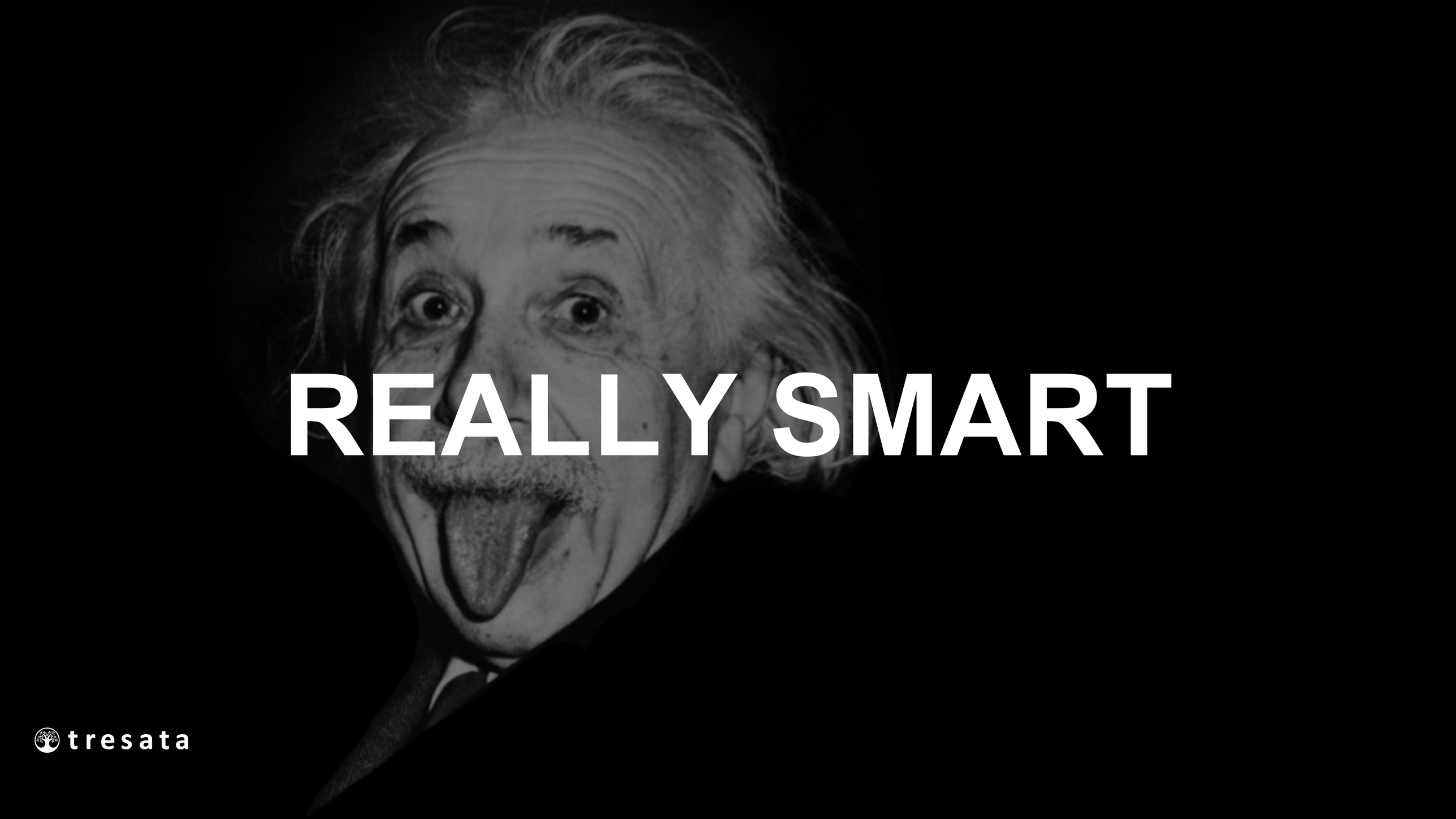
Lack of automation

Inaccuracy


SPARK'ING A REVOLUTION
tresata + databricks™

tresata


REALLY PRECISE
tresata



tresata

Tresata's AML Engine

Our needs for our core graph engine:

- Interactive
  Most queries within few seconds | Large queries no more than 20 seconds
- Graph traversal
  Property graph | Multi-hop | Shortest paths
- Search Engine
  Not just point queries | Ability to become a graph search engine
- No Respect for Size
  >50mm entities | >500mm relations

Tresata's AML Engine

Using Spark to make it happen:

- Interactive
  In memory RDDs | Optimized data structures
- Graph traversal
  Co-partitioning of RDD's | Iteration comes easy | In-house DSL on top of RDD's
- Search Engine
  Dataset Scans | RDD batch processing (in miliseconds)
- No Respect for Size
  Can fit 50MM nodes/ 500MM edges in memory on 10 nodes | 10X bigger if serialized in memory (3X slower though)

QUESTIONS?

www.tresata.com | katie@tresata.com | koert@tresata.com