Déjà vu - when contexts match, opportunities repeat

Kevin B. Pratt, Chief Scientist ZZAlpha Ltd.
Kevin.Pratt@ZZAlpha.com
previously lead in
Advanced Analytics Group at Teradata Aster

Presented at Predictive Analytics World

Las Vegas, NV, USA June 2018

A gallop into the world of big data, artificial intelligence, US equities markets, algorithmic trading and new ways of understanding predictive context linkages.

ZZAlpha LTD.

Consistent returns from machine learning

Kevin Pratt – Chief Scientist and Founder at ZZAlpha

18 years of applied Artificial Intelligence and Advanced Analytics in big data

U.S. AIR FORCE
RINCON RESEARCH
NIH National Institutes of Health
Turning Discovery Into Health
United States Coast Guard
U.S. Department of Homeland Security
Pfizer
HARRIS
DARPA
NCR
TARGET
LOWE'S
DISCOVER
WELLS FARGO
FBI
CHUBB
SAIC
Redefining Ingenuity
TIAA
ICDM
IEEE International Conference on Data Mining
DELL
TERADATA
Aster Analytics
MAYO CLINIC
Sprint
verizon
Predictive Analytics WORLD
VICTORIA'S SECRET
dish
COMPUTER SCIENCE INNOVATIONS, INC.
NORTHROP GRUMMAN
NSA
AT&T
KDD
DIRECTV
REDcard

Every domain cross-pollinates insights.

What is ZZAlpha ?

We use artificial intelligence on big data to identify large stocks that are likely to increase in price over the next week to month.

We publish nightly newsletters with specific recommendations (for over 40 different portfolios).

Formed 2010. Privately held.
Uses objective, public information.
Expertise in artificial intelligence, US markets and big data.

We do NOT sell or promote stocks.

We do NOT make individualized recommendations.

We do NOT receive compensation from anyone except our subscribers.

www.ZZAlpha.com

www.ConservativeStockPicks.com

ZZAlpha LTD.
Consistent returns from machine learning

Take away # 1 Machine Learning works in the stock market

The *learning* algo implemented in 2011 has worked for 7 years unchanged.

www.ZZAlpha.com

www.ConservativeStockPicks.com

ZZAlpha LTD.

Consistent returns from machine learning

Take away # 2
Machine Learning offers scientific benchmarks

We publish ALL our results with benchmarks.

www.ZZAlpha.com

www.ConservativeStockPicks.com

ZZAlpha LTD.
Consistent returns from machine learning

Take away # 3
Machine Learning validates intuitions

A key insight:
A prediction has an optimal duration after which its validity decays.

www.ZZAlpha.com

www.ConservativeStockPicks.com

ZZAlpha LTD.
Consistent returns from machine learning

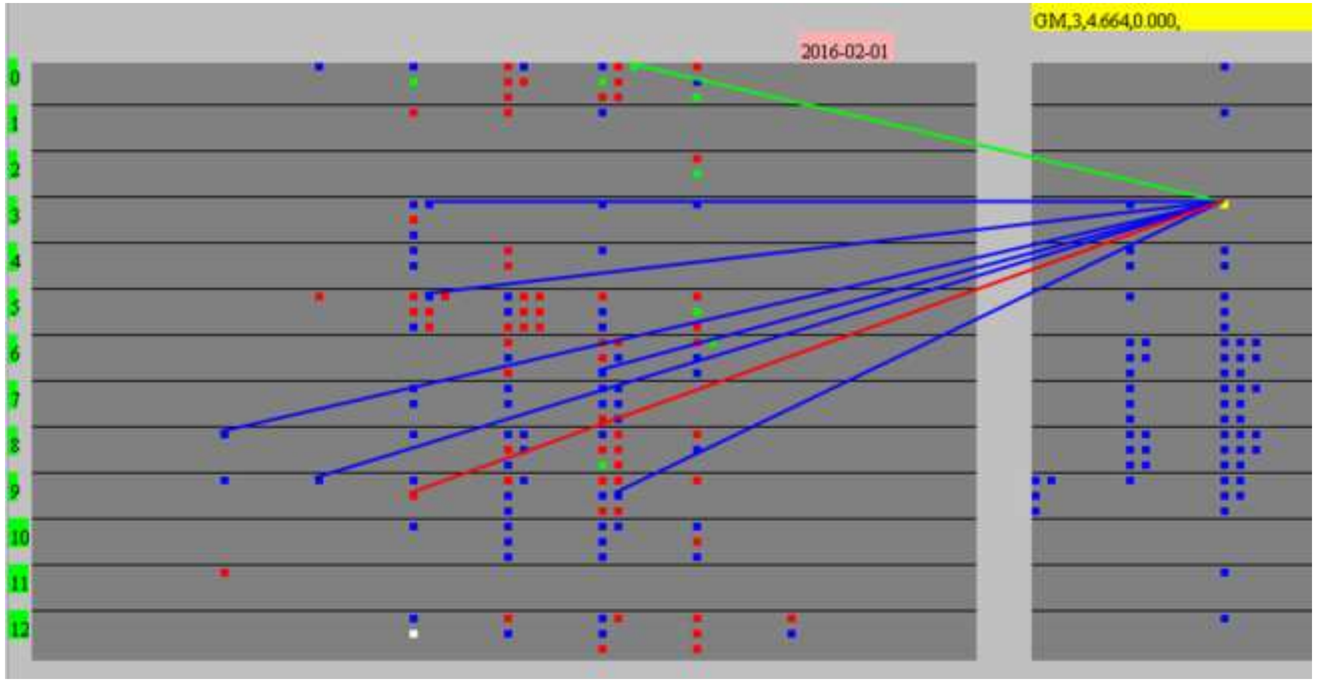
Take away # 4
Machine Learning needs big data

A foundation:
Predictions in complex, dynamic systems require *more* data.

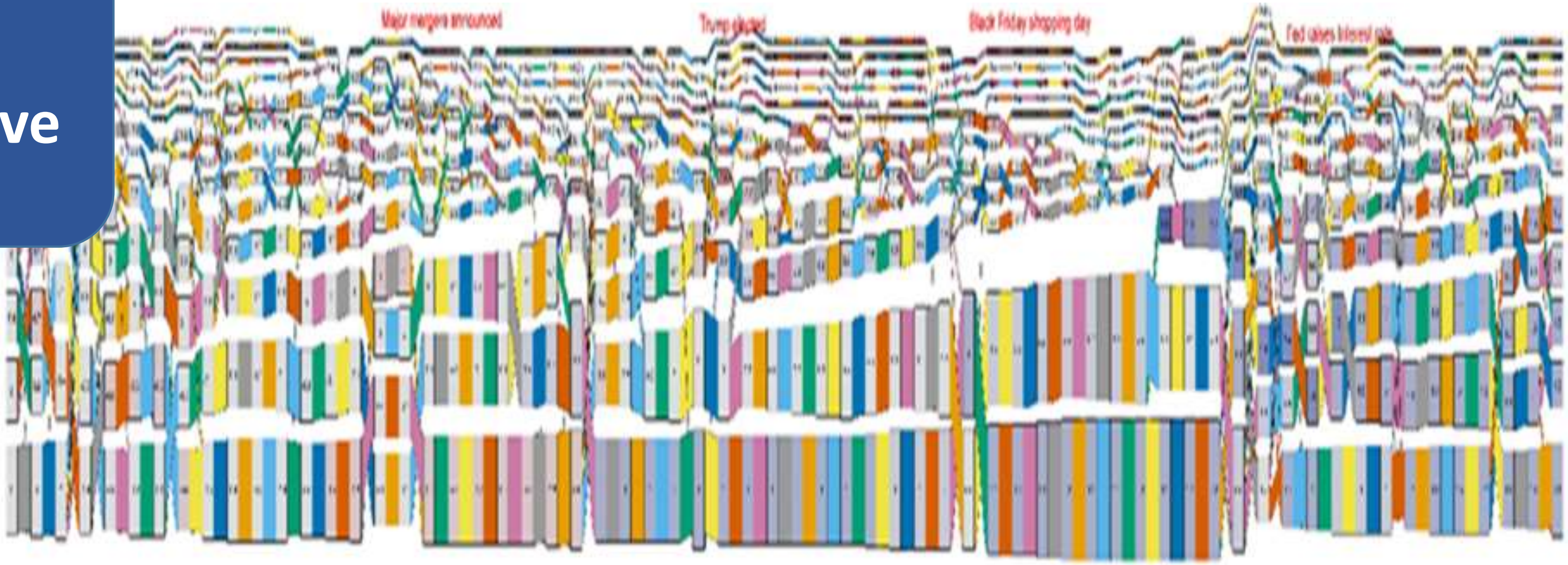
www.ZZAlpha.com

www.ConservativeStockPicks.com

ZZAlpha LTD.
Consistent returns from machine learning



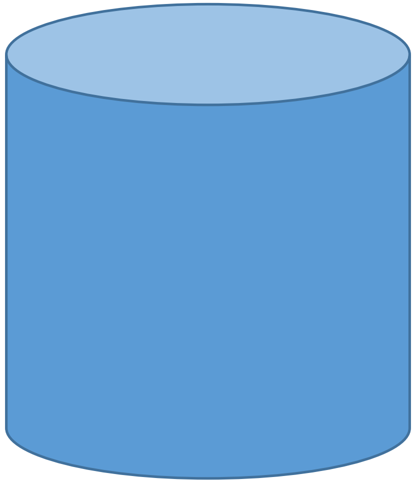
Take away # 5
Temporal visualizations drive understanding



agenda

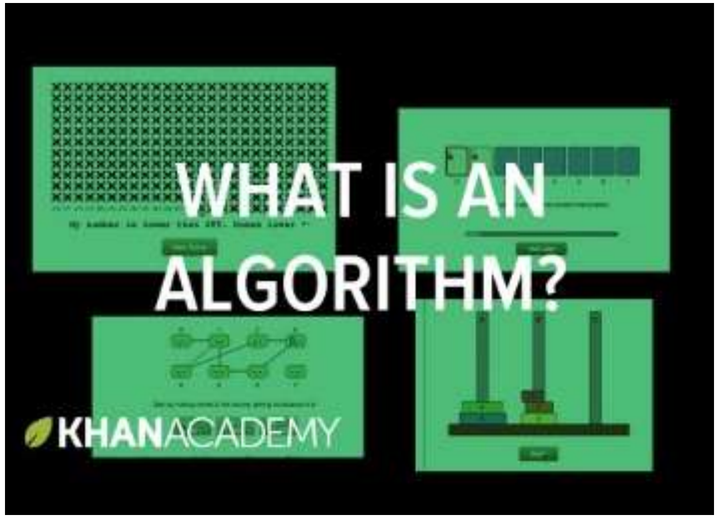
1. Data science and algos
2. Stock Market
3. Puppy learning
4. Evaluation
5. Visualizing dynamic complexity

NEXT: a *VERY* quick look at . . .



Data

Models



Algorithms

What do you mean by "data" ? What's a simple stock example?

Data (garbage in, garbage out)

Price

Orders

Trade

Volume

Aggregates tic

Aggregates minute

Aggregates day

Multiple Exchanges

All data sources have errors, delays, and omissions.

Hindsight data is always better quality than real-time data.

Corporate events – Earnings, Mergers, IPOs, SEC filings, Some news

Analysts' Reports

News Journals

Press Releases

Bloggers and Brokers and Advisors

Popular sentiment and fake news

What do you mean by a “model” ? What’s a simple example?

A model (a line, sometimes called a “trend”)

Traders’ Wisdom:
“A trend is our friend”

A model (a line, sometimes called “dubious”)

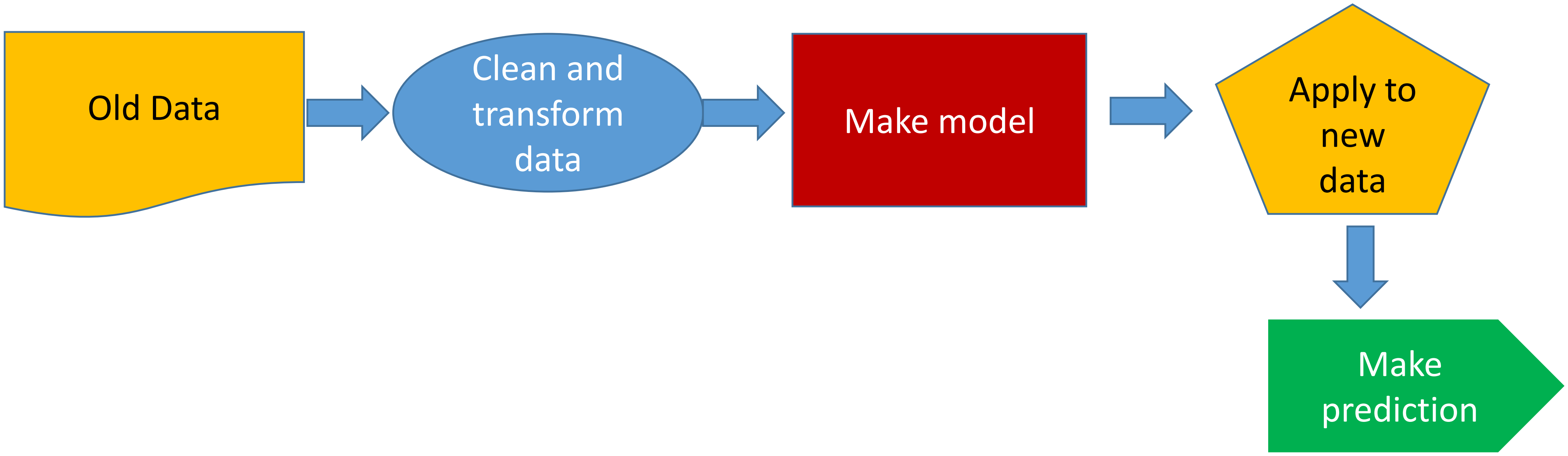
A model (a line, sometimes called “useless”)

Jargon: “Non-Stationarity” is a significant divergence from prior values in a time-series

Non-Stationarities are difficult events to predict with algorithms.

What do you mean by an “algorithm” ? What’s a simple example?

Def: A step-by-step procedure to turn data into a model that you can use to make a prediction.



That’s all ? What’s hard?

There are **6 main types** of machine learning algorithms and maybe **100 flavors** of each and maybe **5000 academic papers** discussing each flavor.

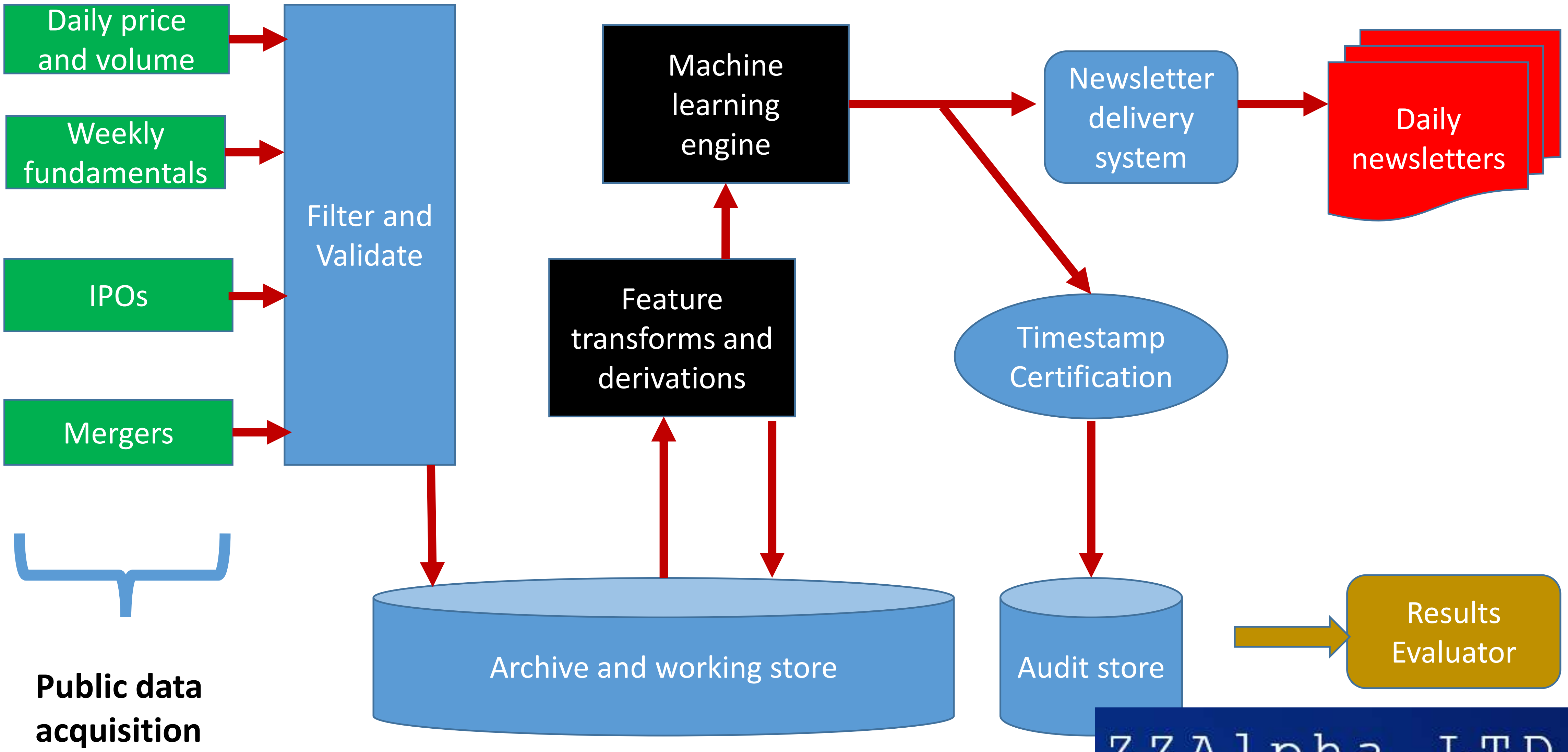
.

1. Linear
2. Tree
3. Probabilistic
4. Network
5. Cluster
6. Flow

I apologize if I left out your favorite.

Artificial intelligence algorithms in production systems

Components of ZZAlpha Learning Architecture in the Cloud



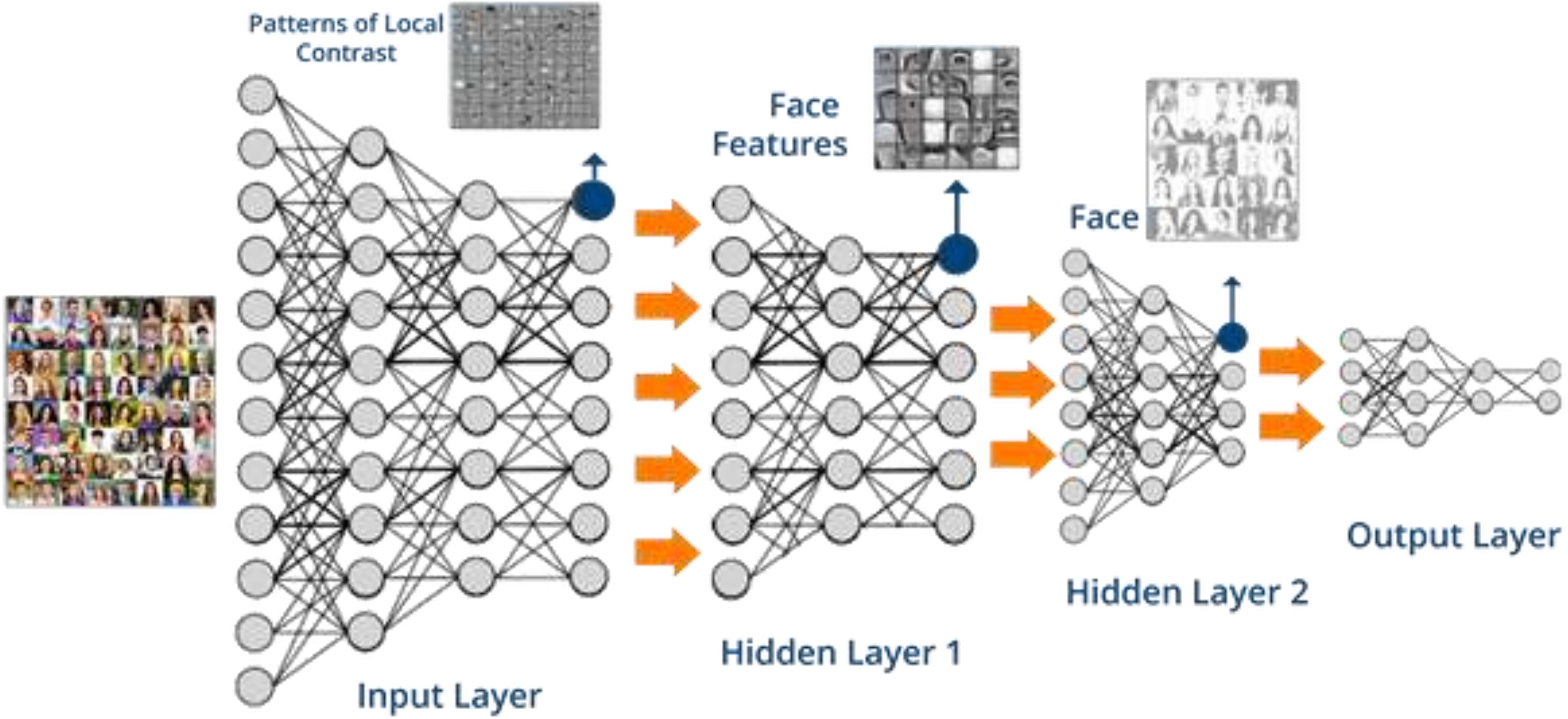
The Dirty Secret of Artificial Intelligence

Too Many …

The Dirty Secret of Artificial Intelligence

Knobs, Inputs, Settings, Transforms, Parameters, Choices, Algos, Tunings . . . . . .

The Dirty Secret of "Deep Learning" Neural Nets



Knobs, Layers, Nodes, Activation functions, Meta-layers, Tunings . . .And SLOW learning

The rule for all production systems:

If something can go wrong, it will go wrong *and* at the worst possible time.

And, **YES**, **this applies to algorithmic trading systems.**

How do I love thee? Let me count the ways . . .
ERROR> out of memory !
Elizabeth Barrett Browning

Will a machine learning algo always get "smarter"?

No. It may ingest and learn from bad data.

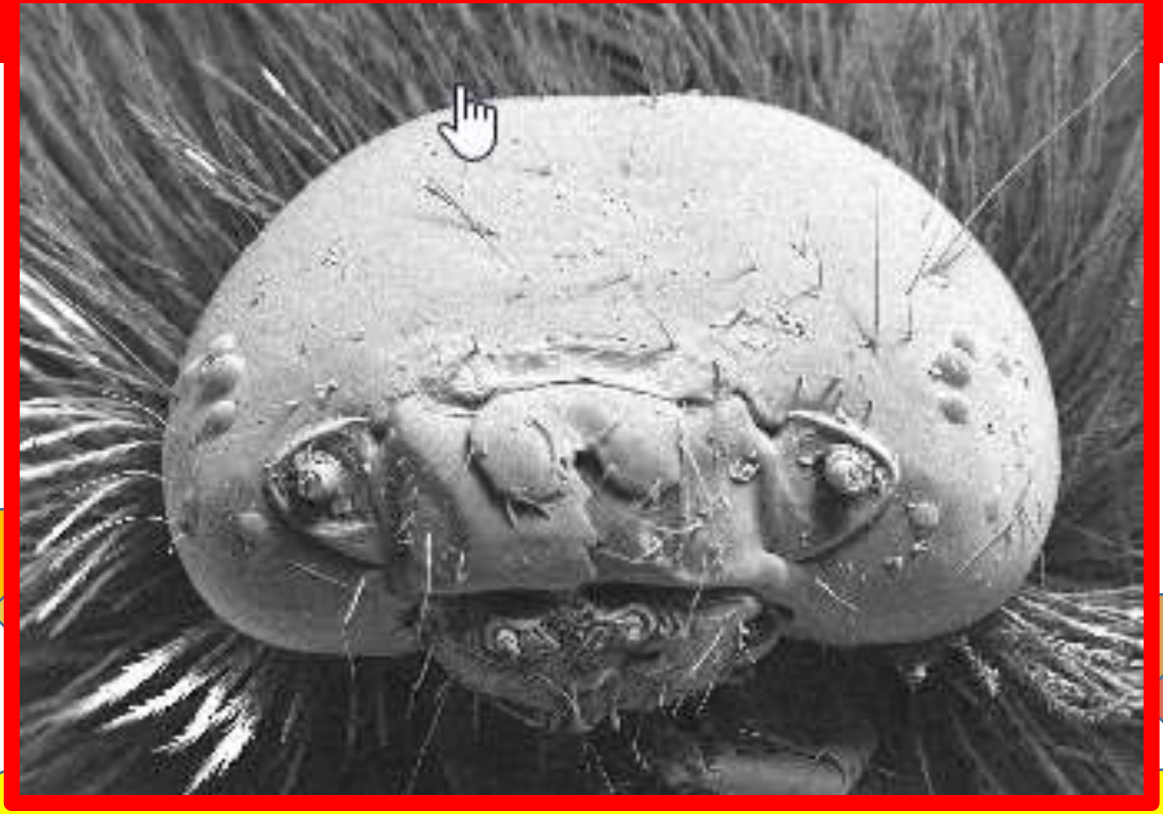
No. It may learn pathological patterns.

No. It may not forget wrong stuff fast enough.

No. Adversaries may learn how to game it.

Will a machine learning algo always get "smarter"?

No. It may ingest and learn from bad data.



No. Adversaries may learn how to game it.

And it may have a weird bug.

And size (data size) matters.

How do you prove it works?

Really. Not in a TV soundbite.

Proof protocol components

Proof Protocol for a Machine Learning Technique Making Longitudinal Predictions in Dynamic Contexts

ACM KDD presentation August 2015 Sidney AU

Kevin B. Pratt, Chief Scientist*

1. *Contemporaneously verifiable discrete predictions*
2. *Deterministic computability of repetitive longitudinal application of predictions*
3. *Imposition of realistic costs and context constraints during evaluation*
4. *Exposure to diverse contexts*
5. *Statistically significant excess benefits relative to* a priori *benchmarks and Monte Carlo trials*
6. *Insignificant decay of excess benefits*
7. *Controlled risk and absence of pathologies*
8. *Extended duration real-time trial "in the wild"*

ZZAlpha LTD.
Consistent returns from machine learning

Good Algorithmic Science

Rule 1:

Measure and measure accurately.

Rule 2:

Measure what matters.

Rule 3:

Document your knobs.

Factoid:
Modern Scientific Method originated ~1020 by Ibn al-Haytham and Abu Rayhan al-Biruni

Good Algorithmic Science

Rule 1:

A Learning Algorithm Can Not Learn Well from Bad Measurements

Well, DUH!

F

F

Document your knobs.

oid:
Scientific
originated
20 by
Haytham and
Abu Rayhan al-Biruni

The Stock Market

Multi-party, Multi-motivation, Multi-timeframe *Adversarial Environment*

Price movements implemented by bots, committees, "fiduciaries" and individuals

Factoids:

75% of trades result from algorithmic trading systems

50% of trades result from high frequency trading (HFT)

54% of assets held by institutions

3% of assets held by hedge funds

34% of assets held by individuals/families

10% of assets held by discretionary investors

Sources: Bloomberg,JPMorgan 2013

Multi-party, Multi-motivation, Multi-timeframe *Adversarial Environment*

Factoids:

Passive ("index") funds account for 29% of the US market.

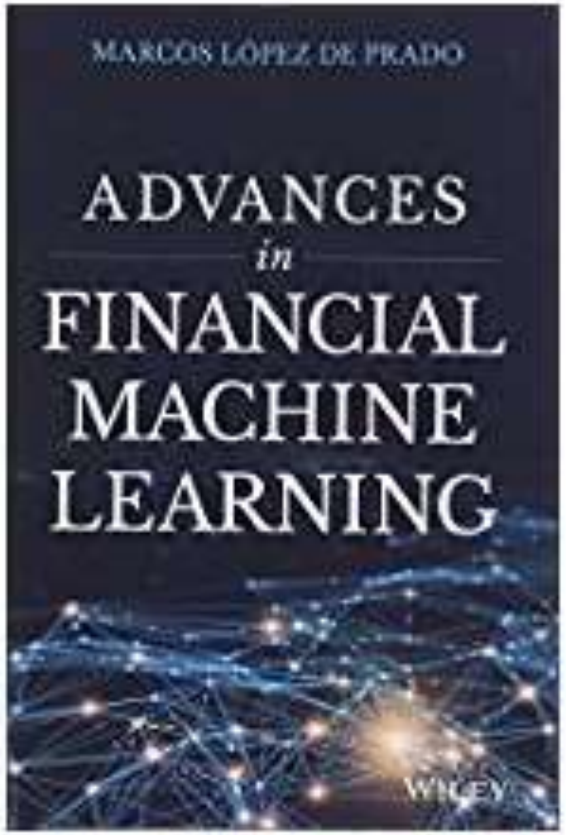
Reuters, Feb. 2, 2017

EVERYBODY ELSE is seeking an "edge" to beat the market.

Algos are a legal source of "edge"

(when using legally obtained data) .

State of the Art 2018



Advances in Financial Machine Learning Feb 21, 2018

by Marcos Lopez de Prado

Hardcover · 30

$33.52 ~~$50.00~~ prime

Get it by **Wednesday, May 30**

FREE Shipping on eligible orders

More Buying Choices

$29.52 (24 used & new offers)

Kindle Edition

$26.39

Get it **TODAY, May 28**

What is the right question?

The fundamental algorithmic trading question:

What are the best choices NOW
for the best NET results
at the END of my time window?

Given:

- My current total “pot” of funds
- My drawdown (“risk”) limits
- Compounding of interim results during the time window

Stock traders ! Are you instrument rated?
(should you be allowed in the cockpit with algorithms)

Déjà vu - when contexts match, opportunities repeat

We also call it “Puppy learning”

ZZAlpha LTD.

Consistent returns from machine learning

What are the most important "factors" (cards) in the context of this winning hand?

Card players live in a world of **set** analysis

A “**set**” is a simple example of a “**network**” (that has **edges** connecting **nodes**)

The combinations of linked nodes in a network can get very large, very fast.

“Network” Analysis?

Hey Dude !
TOO mathy !

"**Network**" Analysis? Nobody much cared until …

Larry Page and Sergey Brin

↓

PageRank Algorithm

↓

Google

↓

$ 765 Billion

Algorithm – “puppy learning”

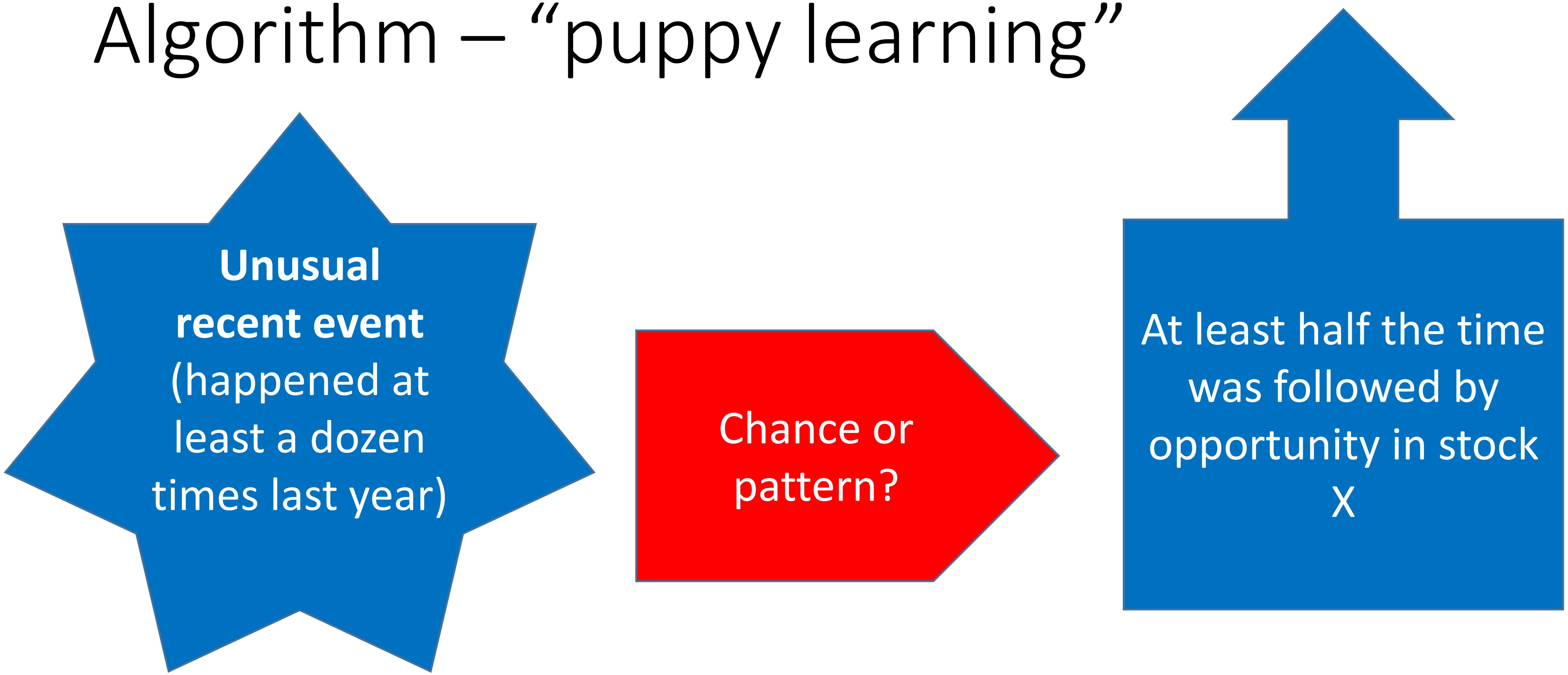
I put on my sox.

Chance or pattern?

Puppy gets a walk.

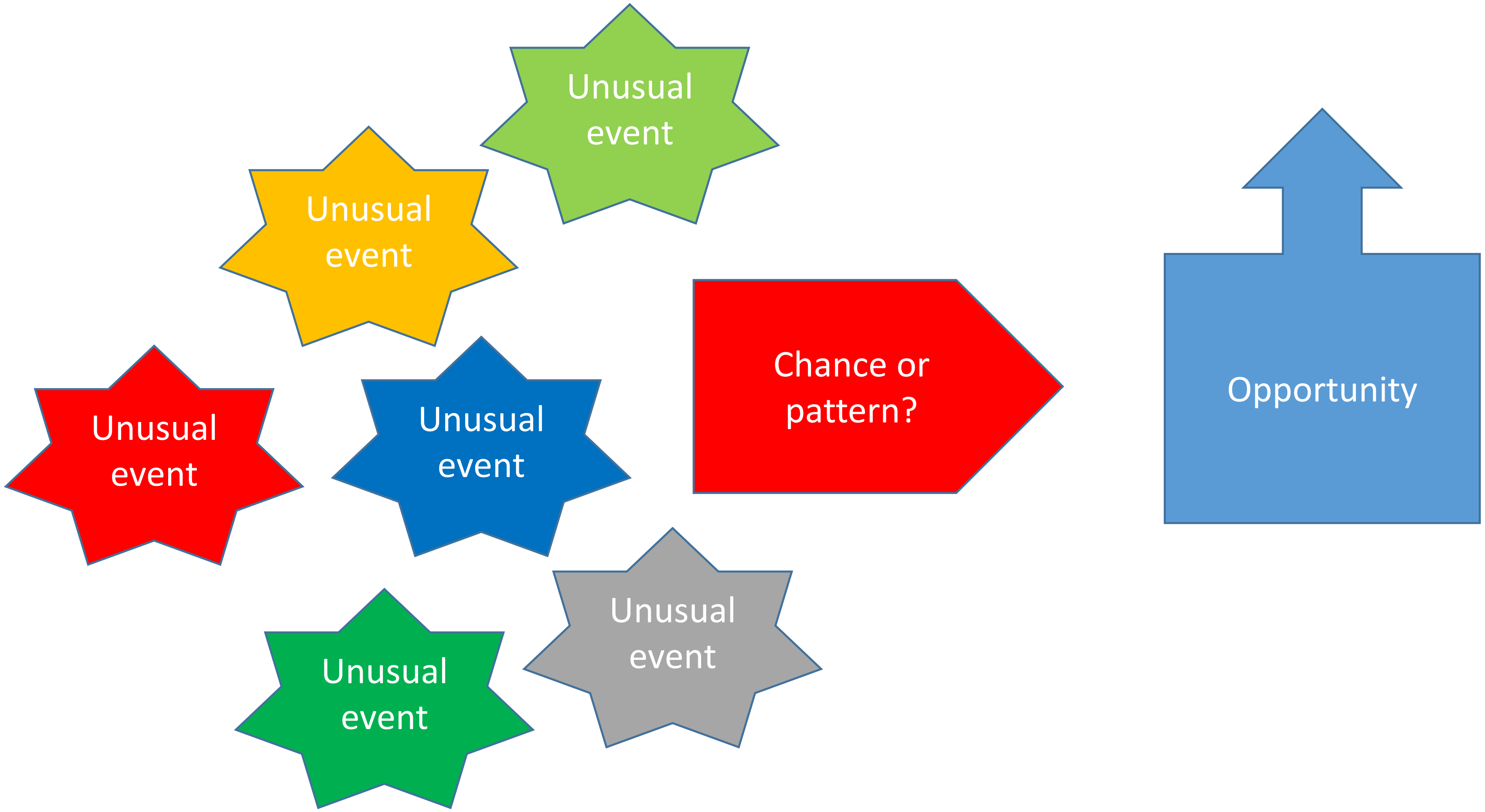
The inspirational puppy

Algorithm – "puppy learning"

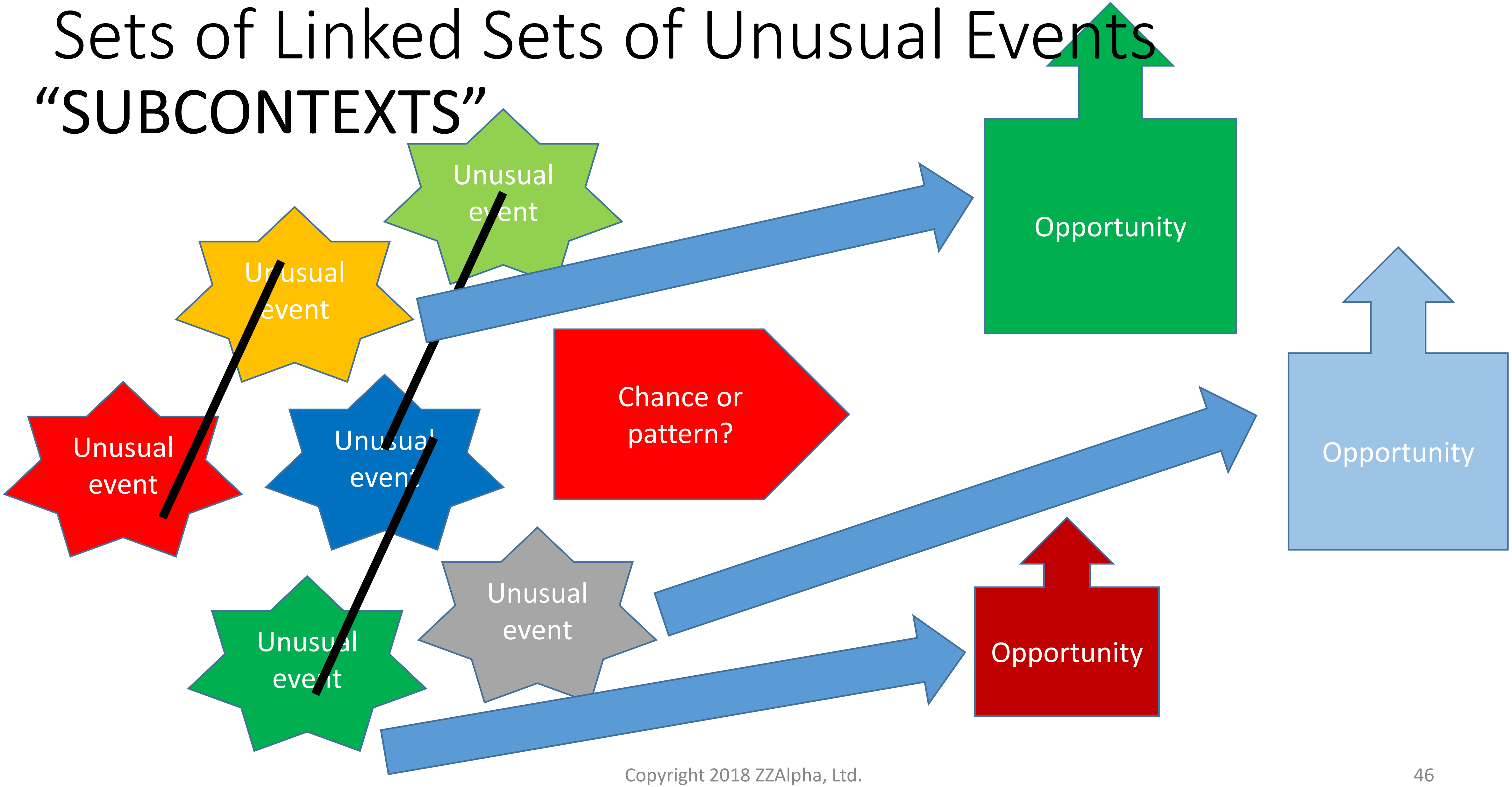


We use multiple meanings for "unusual"

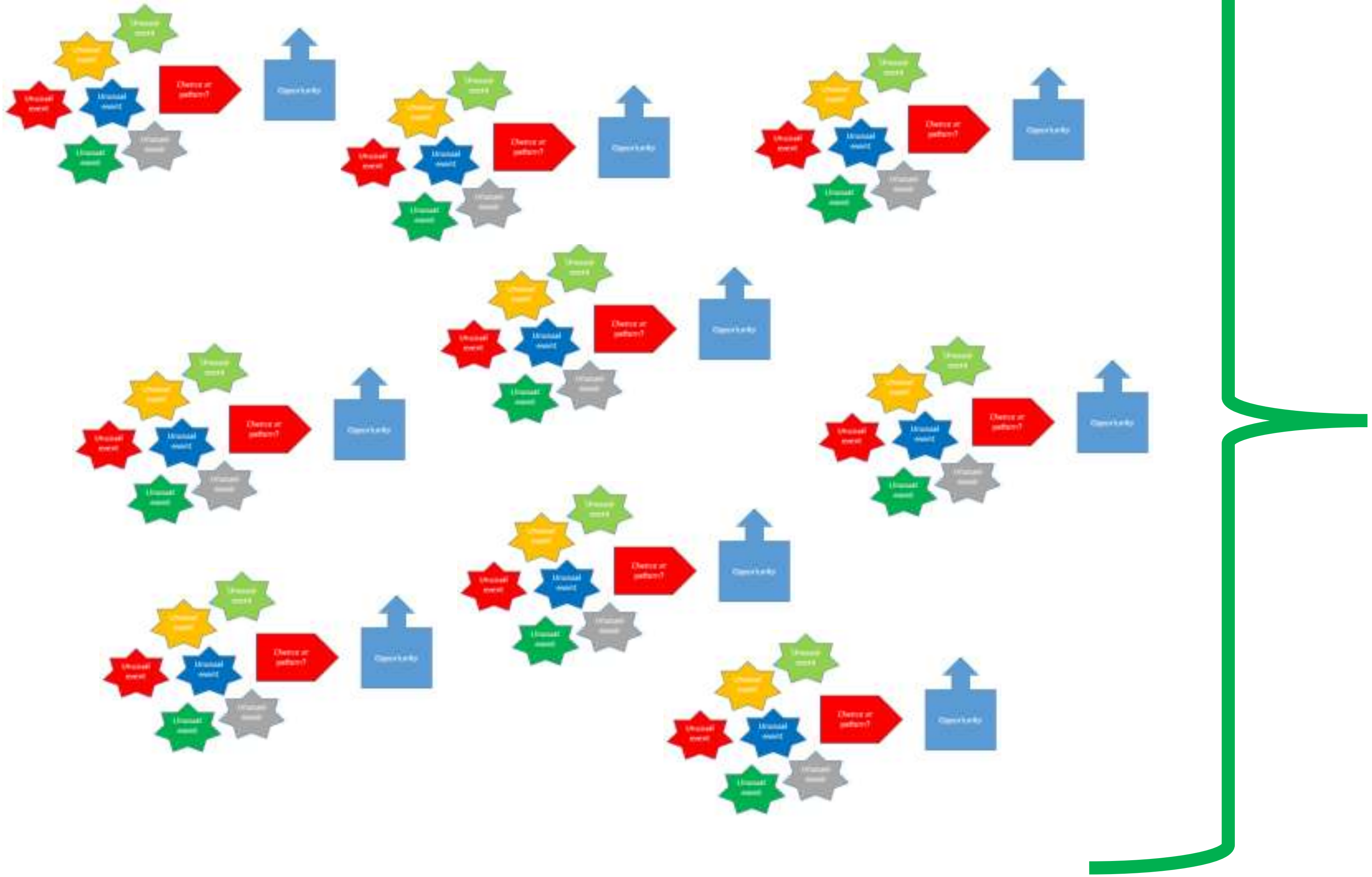
Set of Unusual Events = "CONTEXT"



Sets of Linked Sets of Unusual Events
"SUBCONTEXTS"



Many Sets of Linked Unusual Events: Many Opportunities



800 possible opportunities today

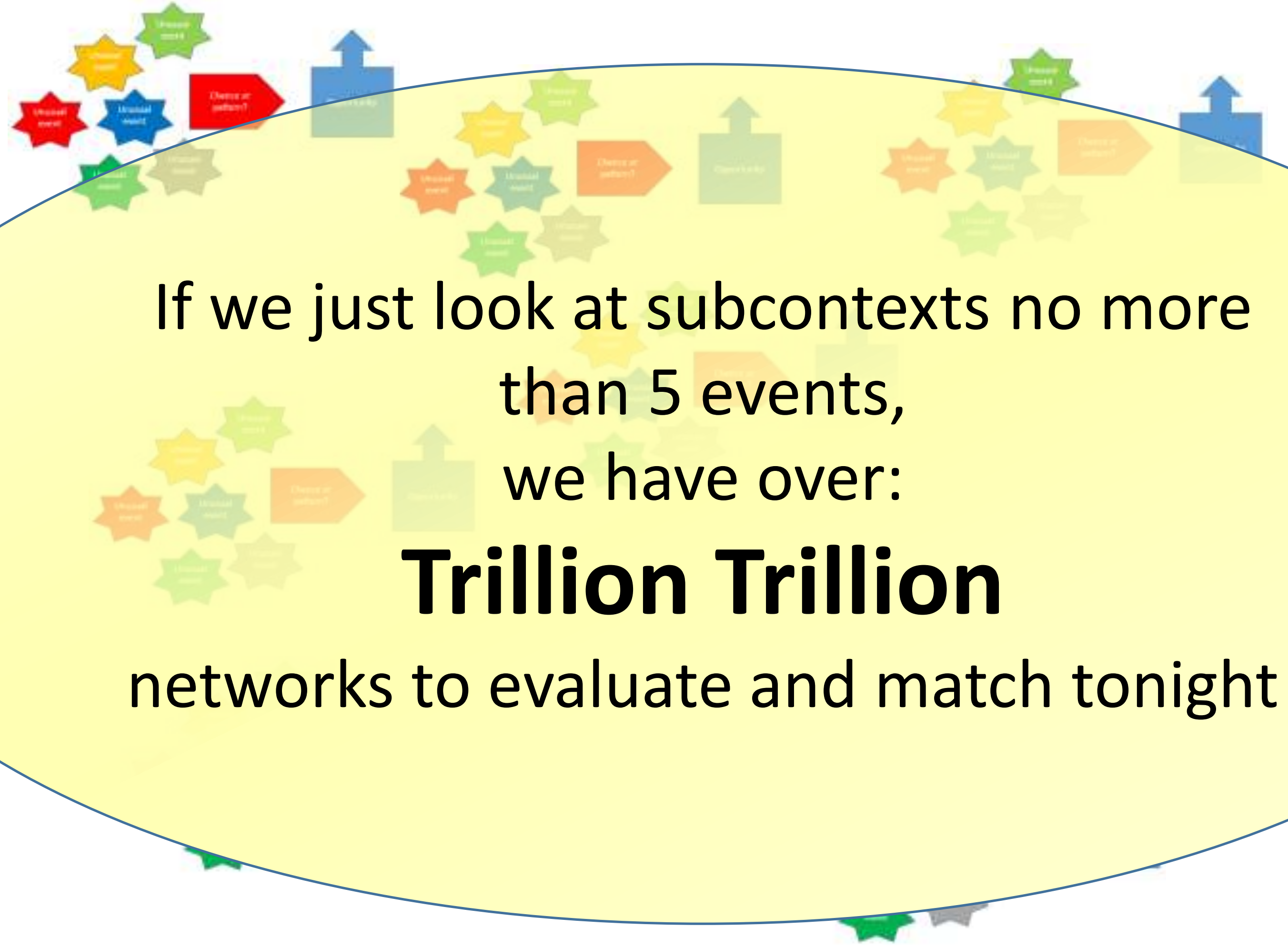
5000 possible precursor events recently

Which opportunity will turn out best **IN TIME HORIZON?**

But, remember the footnote ?

The combinations of linked nodes in a network can get very large, very fast.

Many Sets of Linked Unusual Events: Many Opportunities



If we just look at subcontexts no more than 5 events,
we have over:

Trillion Trillion

networks to evaluate and match tonight

800 possible opportunities today

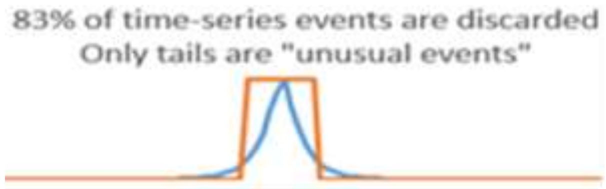
5000 possible precursor events recently

Which opportunity will turn out best **IN TIME HORIZON?**

Innovations for implementation

(inspired by the puppy):

1. Discard all the "ordinary events" and just look at the unusual ones. *(Throw away 83% of the raw data!)*



2. Instead of looking **along one** (or a few) time series to try to predict, look **ACROSS all** of them.

3. Evaluate links for trustworthiness before accepting them for reuse.

These make everything much easier to compute.

Puppy Learning Algo Evaluation

When we do it, is it better?

Conservative standards for algo evaluation **from the perspective of an investor:**

After applying **realistic trading costs** (commissions, spread) and
Looking at **1 year, 3 year and 5 year results:**

1. Is it **better than** buy and hold **SPY** (S&P 500 tracking ETF) ?
2. Is it **better than** an appropriate **benchmark** ETF (that includes dividend reinvestment)?
3. Is it **better than** an **unweighted** random selection from its **population**?
4. When it recommends both long and short, are **anti-long results much less** than long results?
5. Is it **better across diverse contexts**?

Is it better?

Conservative standards for algo evaluation **from the perspective of an investor:**

After ap
Looking

1. Is it
2. Is it end
   reinv
3. Is it on?
4. Whe **ch less**
   than
5. Is it **better across diverse contexts**?

Conservative investors ask:

Would you

bet your inheritance *and*

your first born child

on it?

Is it better?

Example of recommendations from S&P500 stocks

Timeframe:	5 yr	3 yr	1 year (2017)
Benchmark: **SPY** (S&P 500 ETF)	**15.4**	**10.4**	**18.8**
Unweighted Sector (496 stocks):	15.1	9.0	15.5

Annualized returns. SPY includes dividend reinvestment.

Is it better?

Timeframe:	5 yr	3 yr	1 year (2017)
Benchmark: **SPY**	**15.4**	**10.4**	**18.8**
Unweighted Sector (496 stocks):	15.1	9.0	15.5
ZZAlpha (2011 Algo)	24.0	15.1	13.3
2016 Puppy Learning	28.7	21.6	12.6
Anti-Long	-4.2	-4.5	-18.0

26.5
BigCap 100
2017

Example from one of over 40 portfolios. Results vary. See ZZAlpha.com for detailed results.

Algo annualized returns. Assumes $10 round-trip trade commission. 5(20) day hold of 5 recommendations per newsletter. Compounded results. SPY includes dividend reinvestment.

How do we visualize what is happening over time?

How do we locate pathologies in the data and algo?

A bi-partite graph

5000 left x 100 right nodes

(no links shown)

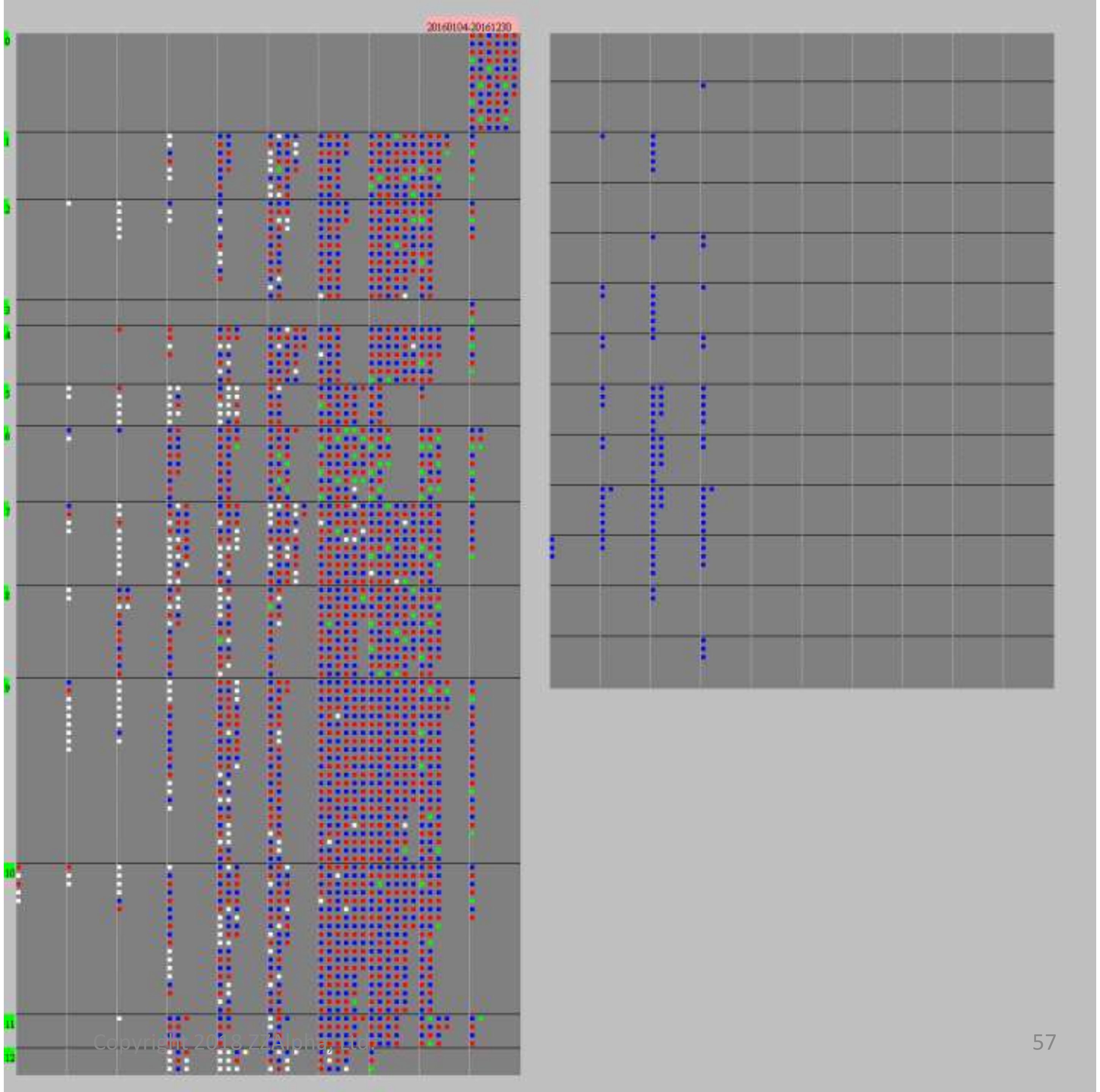
Left= all unusual event types occurring in 2016

Right= unusual price rises (in the largest 100 stocks) in 2016

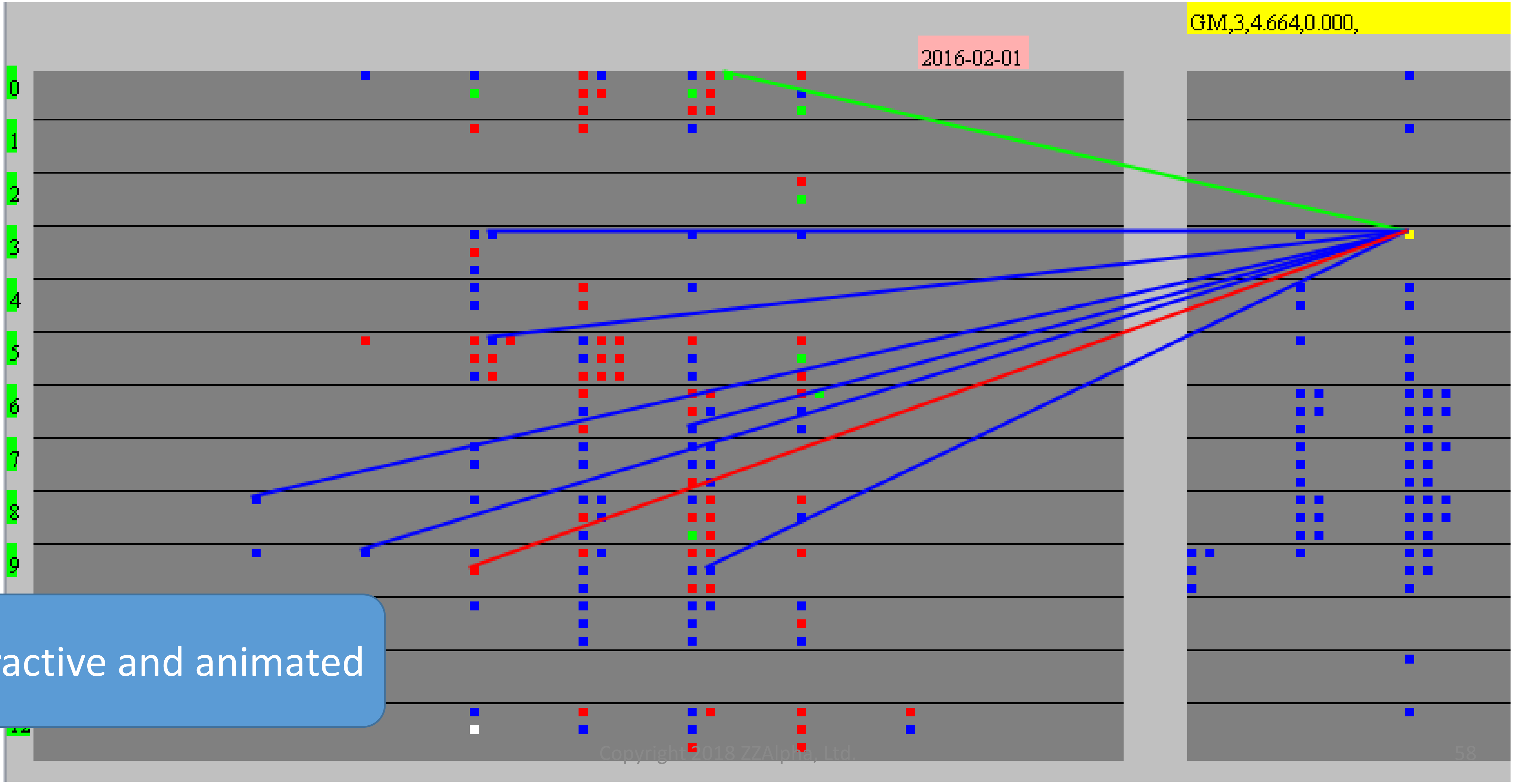
Node color = an event type group (e.g. “peak”)

Horizontal groups= economic sectors

Vertical groups = other metadata groups

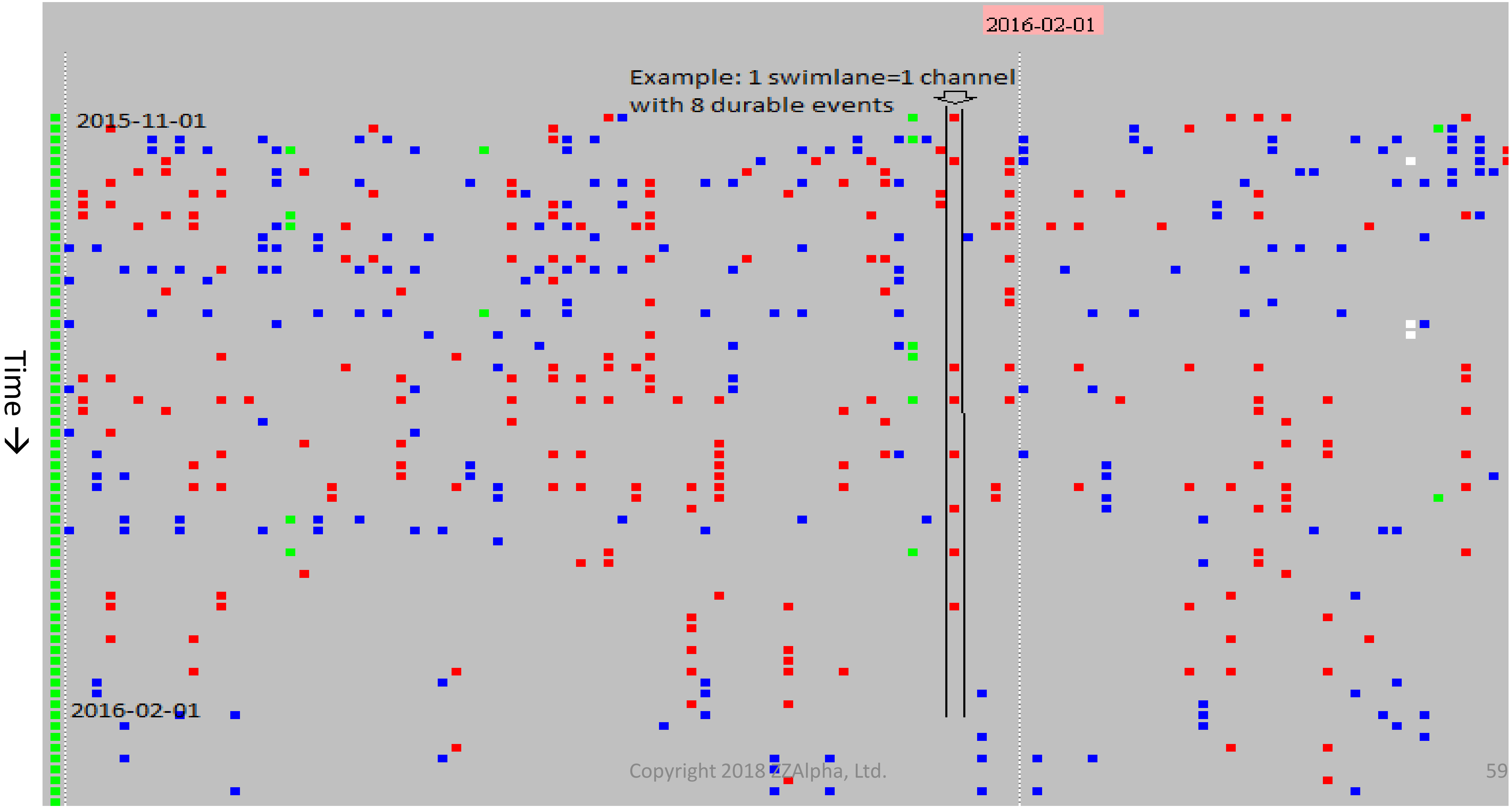


Bi-partite graph of one set of **unusual events (on left) linking to an opportunity** (on the right, to buy GM) on Feb 1, 2016

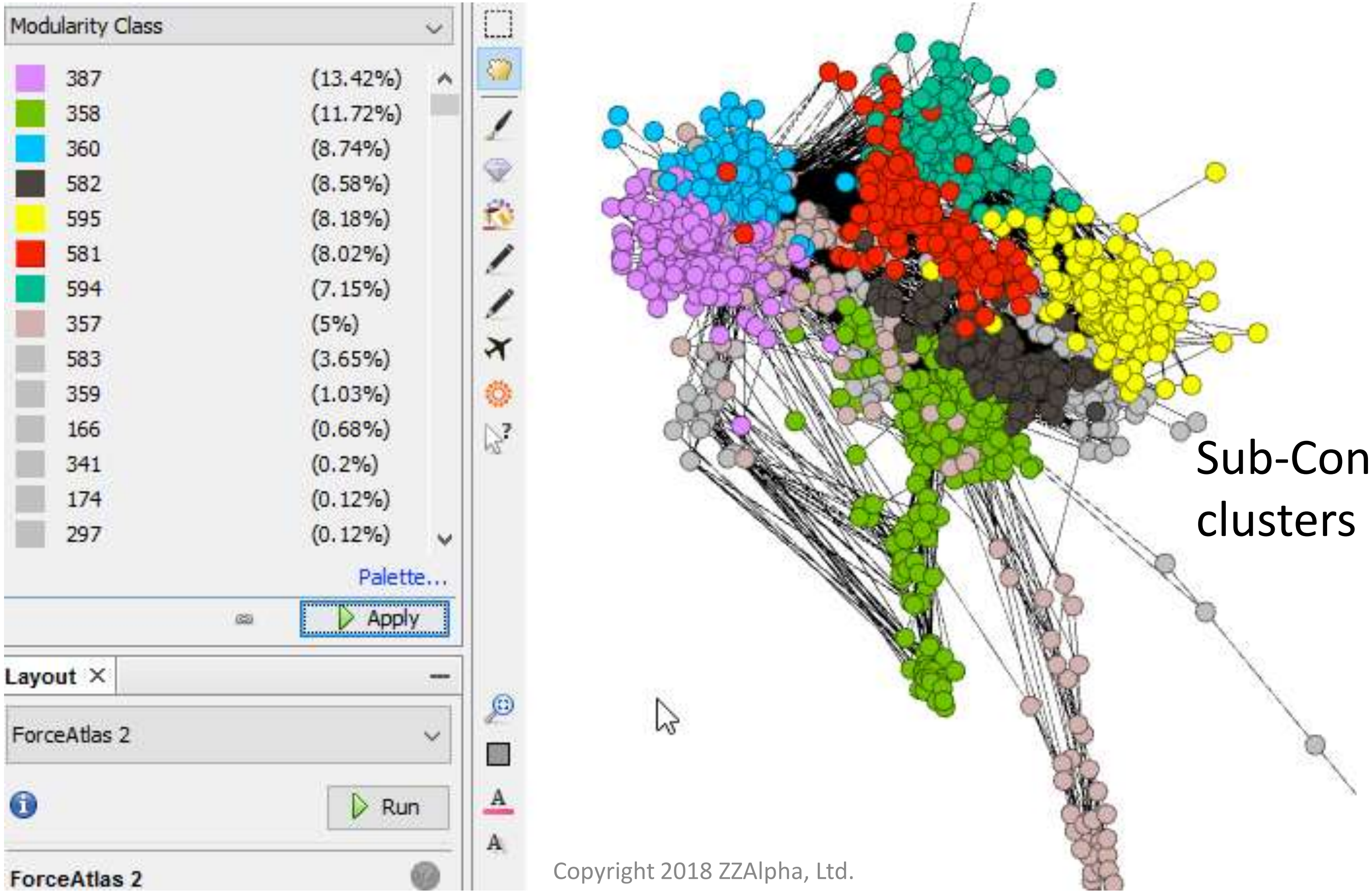


Interactive and animated

Historic distribution of unusual events (each colored box) linked to GM over each day of the prior 3 months

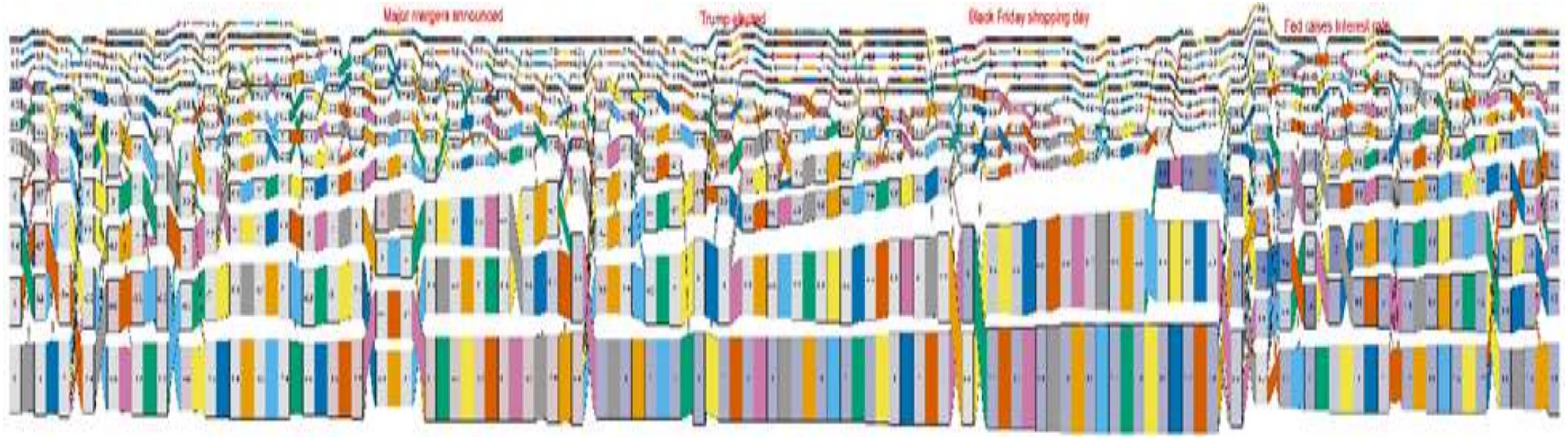


Snapshots of clusters of unusual events on a day are less useful than understanding **evolution of those clusters** (next slide)



Sub-Context clusters

Evolution of clusters of sub-contexts over 6 months showing **merges and splits** and average cluster results



To help answer the question: how do external events impact changes in linked sets of unusual events?

In closing

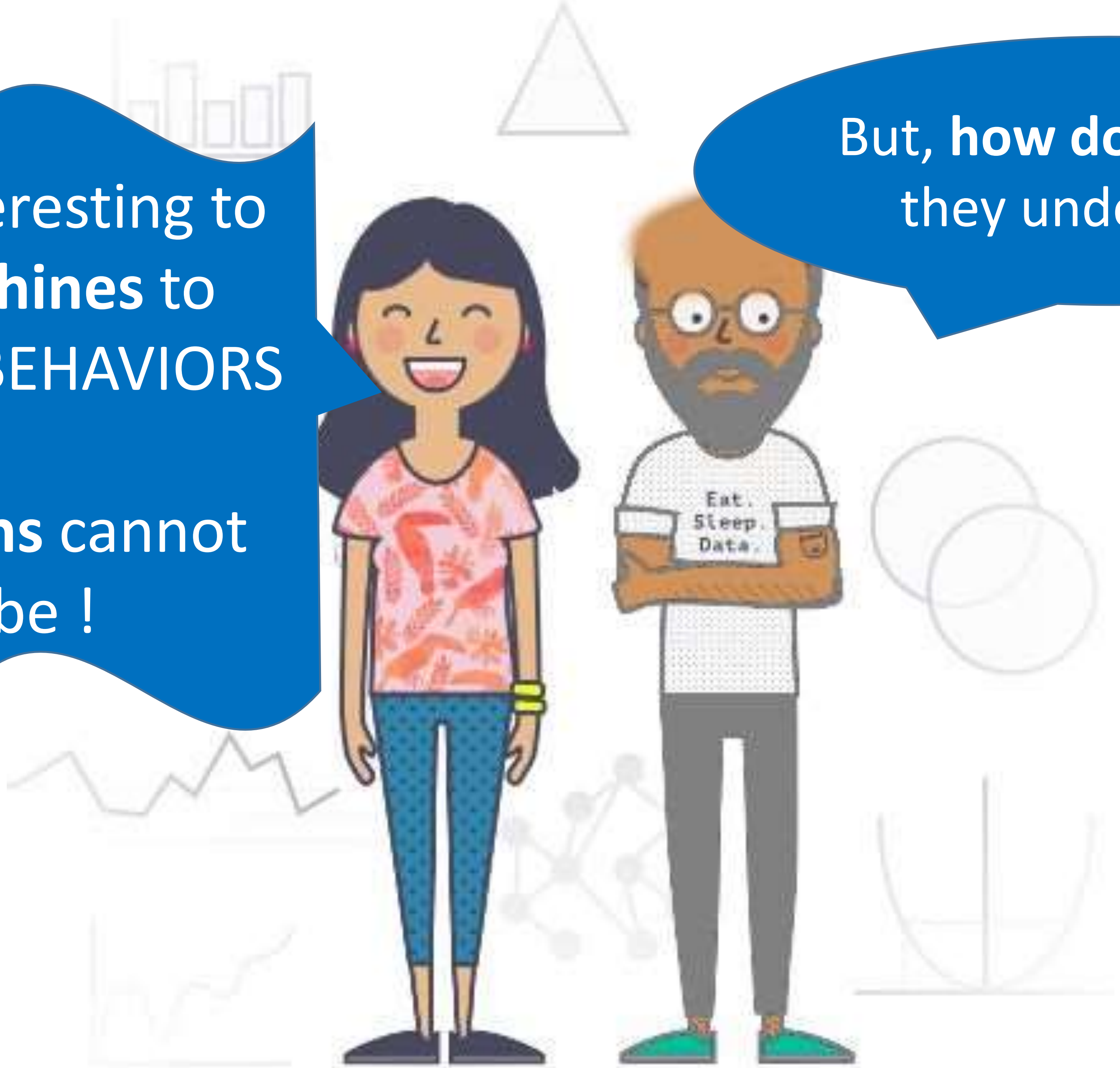
One final hint

Beware the Dangerous Data Scientist

The fool
with a tool

It is very interesting to **teach machines** to understand BEHAVIORS

that **humans** cannot describe !



But, **how do you know** they understood?

It is very interesting to **teach** HORSES to understand BEHAVIORS

that **humans** cann

describe !

But, **how do you know** they understood?

So, you'd rather ride on a driverless car?

Thank you.

Anxiously welcoming your questions . . .

Kevin.Pratt@ZZAlpha.com

www.ZZAlpha.com