

RSA® Conference 2015

San Francisco | April 20-24 | Moscone Center

SESSION ID: DSP-W01

MORE APT THAN YOU THINK: DATA PROTECTION AT MASSIVE SCALE

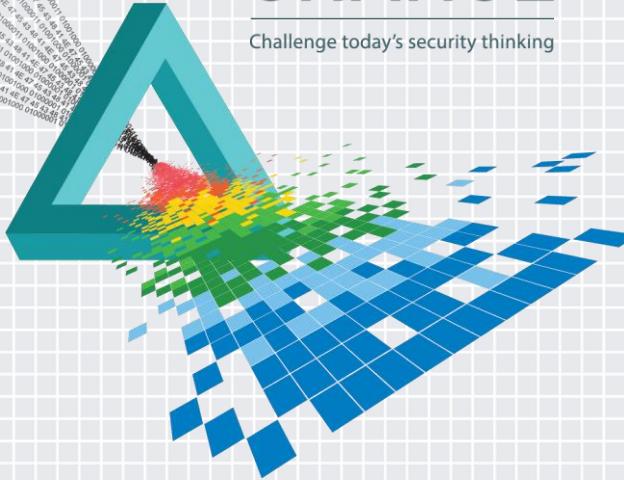
Davi Ottenheimer

Senior Director of Trust, EMC
@daviottenheimer

Wednesday, Apr 22nd 8:00 AM | West 3006

CHANGE

Challenge today's security thinking



COMMODITY PLATFORMS

AI ML
ANALYTICS

INFO- SHARING

"THINGS"

IT CAME FROM THE MARKET...

RAWR



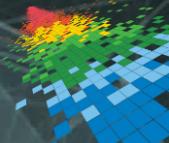
ESSENTIAL DATA NAVIGATION TOOLS

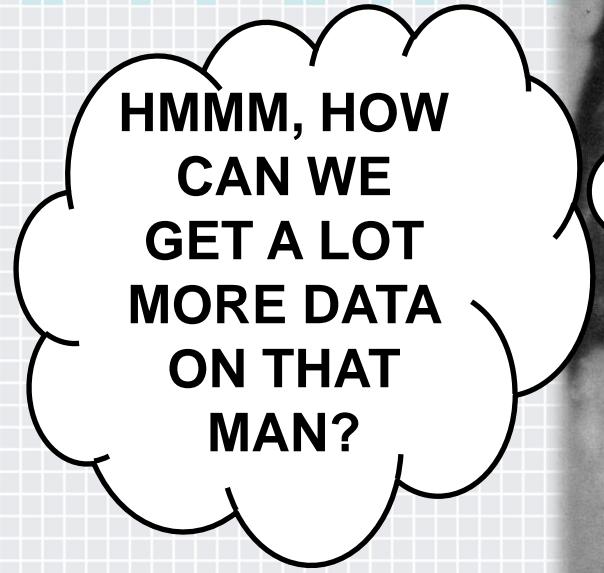
Philosophy

1. Ask Hard Questions
2. Use Correct Reason
3. Clarify Values
4. Find Foundations

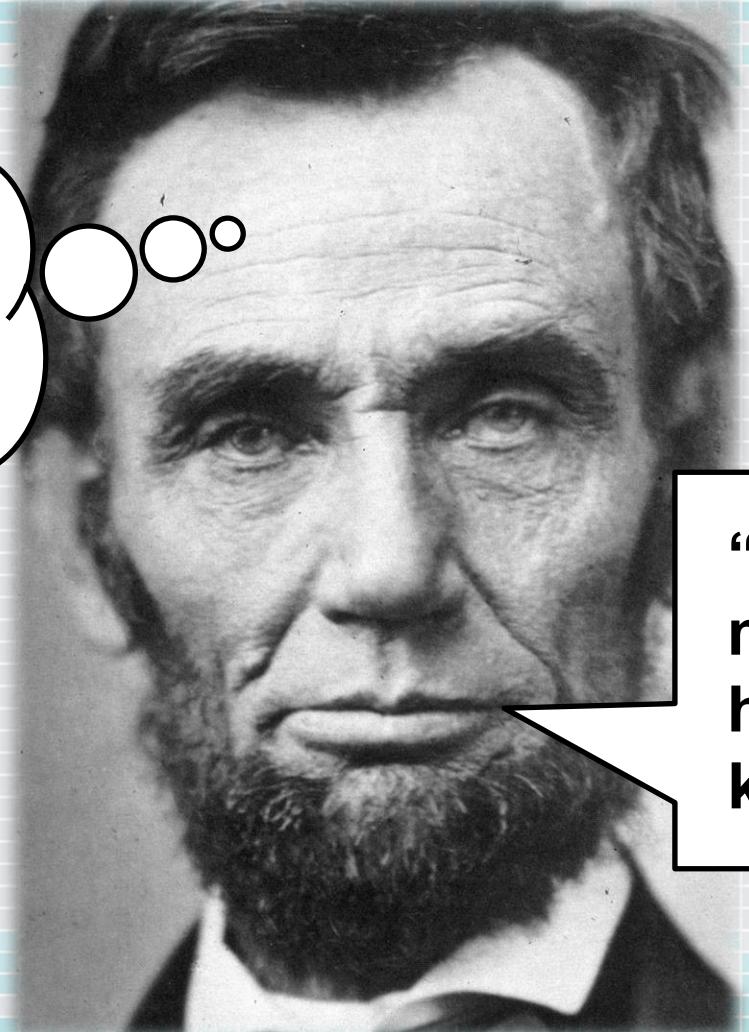
History

1. Identify Sources
2. Prepare/Process Information
3. Analyze to Improve Present Understanding

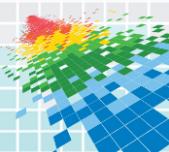


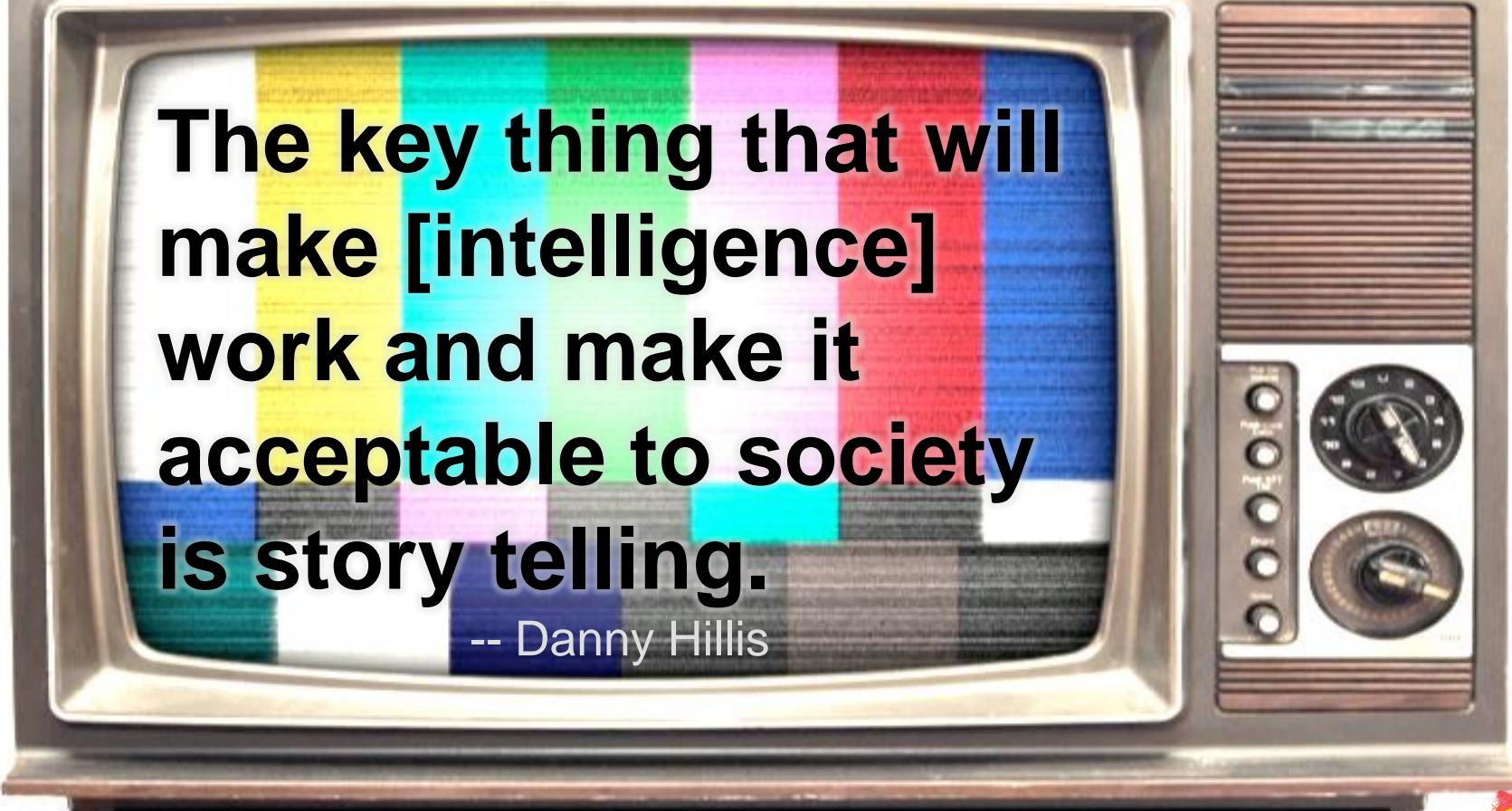


HMM, HOW
CAN WE
GET A LOT
MORE DATA
ON THAT
MAN?



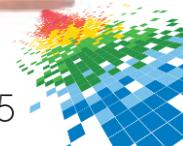
“I don't like that
man. I'm going to
have to get to
know him better.”





The key thing that will make [intelligence] work and make it acceptable to society is story telling.

-- Danny Hillis



112 MILLION FINGERPRINT FILES

1943 FBI*



* 2015 “long-suspected problems with subjective, pattern-based forensic techniques”

~80% US Population

183,950 CARS CHECKED IN 158 DAYS

1965 US CITY POLICE

INSTANT CRIME STATISTICS

The city of Chicago is a good example. From beat to beat all policemen have access to daily crime statistics including the plate numbers on all stolen cars. All files are immediately accessible from an IBM 1410 computer.

The computer also serves as a "home" for records on 90,000 persons and vehicles. Policemen in Chicago reportedly phone or radio questions to their private data bank at a rate of 1400 per day. The reply they receive (average 2 minutes) can mean a lot to a policeman—maybe his life. From such information patrolmen can categorize their suspect in a range that runs from "shoplifter" to "potential murderer."

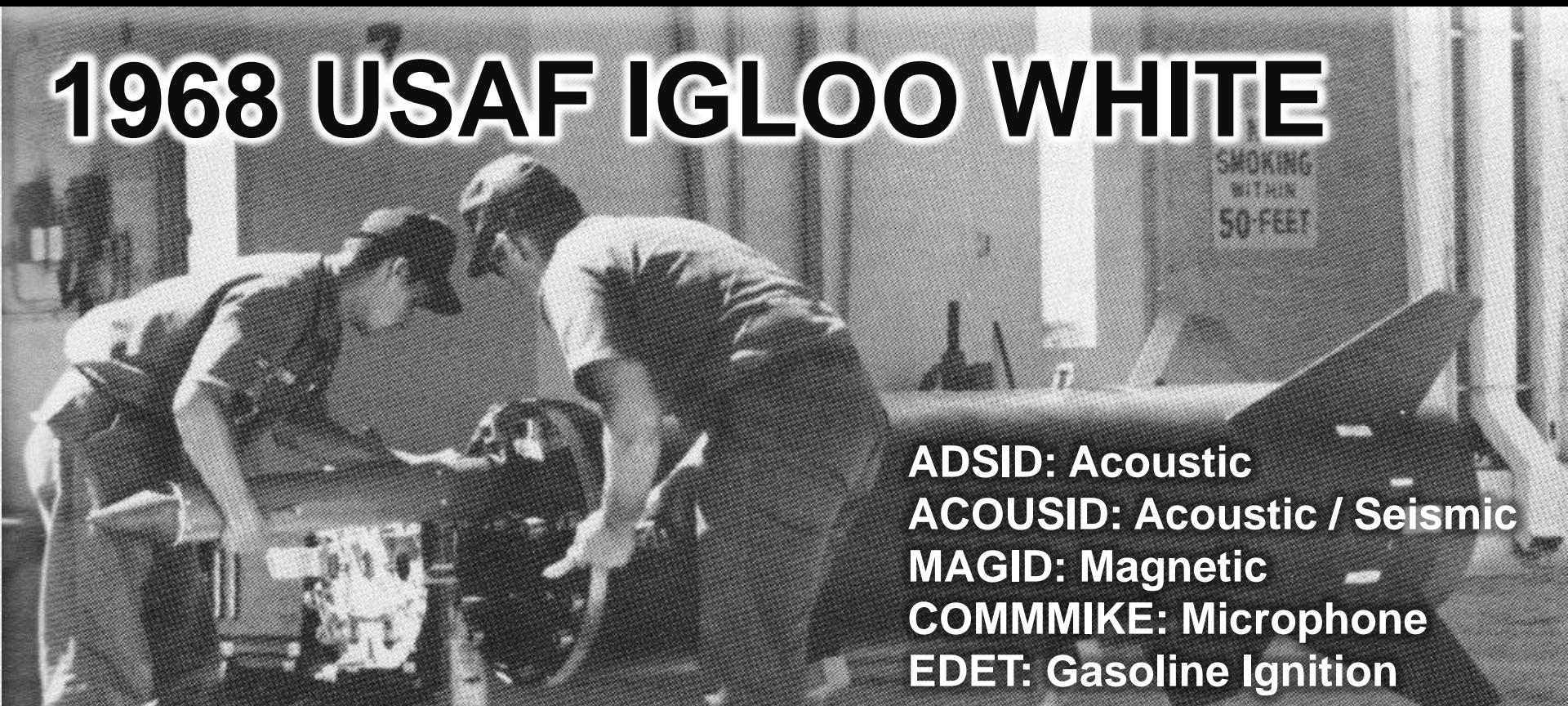
And Chicago isn't the only city that is adding computers to its recruiting list. The results of a recent experiment called "Operation Corral" conducted by the New York City Police Department are equally surprising. Using a Univac real-time computer, "Operation Corral" checked on 183,950 cars in 158 days. Of these, 2932 were either wanted in alarms for stolen cars or plates, or on warrants as scofflaws. A total of 165 people were arrested, many of whom turned out to be wanted on more serious charges ranging from narcotics violations to grand larceny.

http://archive.computerhistory.org/resources/text/Computer_Usage_Company/070910_download/computerusagecompany.computers_and_law_1967/computerusagecompany.computers_and_law_1967.pdf

"whatever the solution, it will be enhanced by computer usage"

20,000 SENSORS DROPPED INTO LAOS

1968 USAF IGLOO WHITE



ADSID: Acoustic

ACOUSID: Acoustic / Seismic

MAGID: Magnetic

COMMIDIKE: Microphone

EDET: Gasoline Ignition

Operational Cost \$1 Billion/Year

**FAST
FORWARD
TO TODAY...**

MATH MODELS = IMPERFECT STORIES

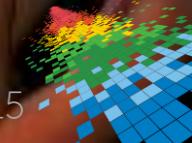
```
01100010  
01100001  
01100010  
01111001  
01100010  
01111001  
01100010  
01111001
```



MATH MODELS = IMPERFECT STORIES

A photograph of a group of people gathered around a glowing digital cube. The cube is illuminated from within, casting a bright light. The people are looking intently at the cube, some holding smartphones to capture the moment. The scene is set in a dark room, with the glow of the cube being the primary light source.

01100010 01100001 01100001 01100010 01100001
01100010 01111001 01100010 01100010 01111001
01100010 01111001 01100010 01100010 01111001
01100010 01100001 01100010 01100010 01111001
01111001 01100010 01100001 01100010 01111001



2015 DRONE PLATFORM FOR...



"I see now why they recommended *margarine* for error"

880M Hectares of Cropland

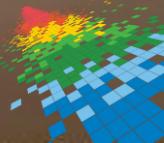
100K Mines & Quarries

300K Wind Turbines

4M Cell Towers

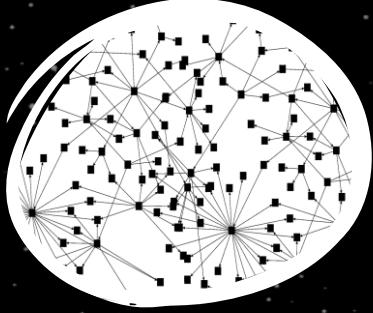
600K Bridges

...





+



+



CONTEXT

Persuasion, Empathy,
Problem Solving,
Negotiation, Physical Skill

THREAT

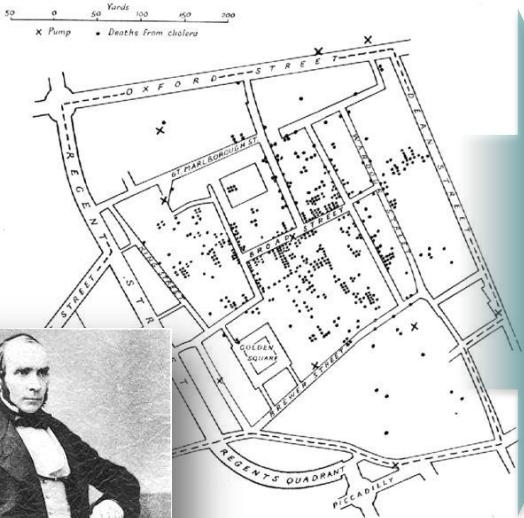
Familiarity, Change,
Persistence

CAPABILITY

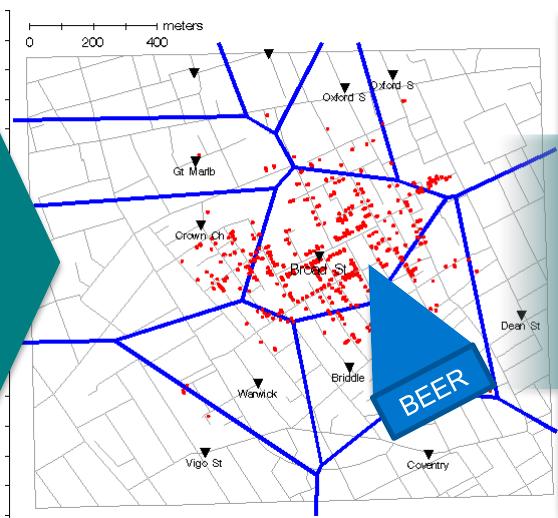
Volume, Velocity,
Variety,
Vulnerability

2012 RSAC: DATA BREACHES AND CHOLERA

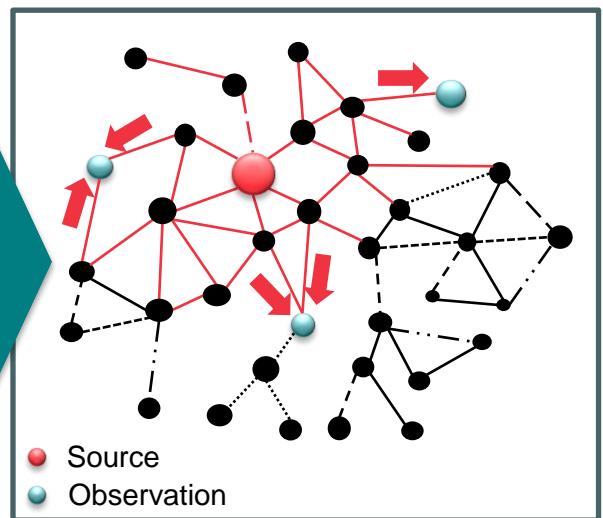
1854: London Ghost Map



1854: London Death Polygons



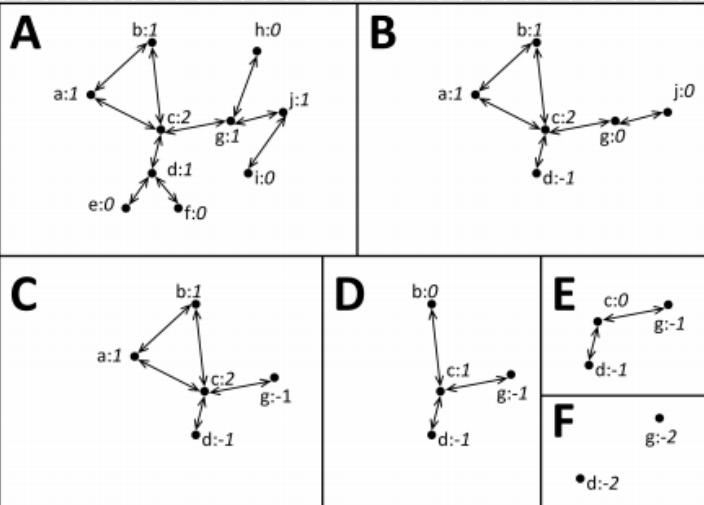
2012: Breach Investigations



(2010 Rinderpest, 2013 AIDS)

Who was trusted to correctly attribute threat? **Who** was authorized to take action?

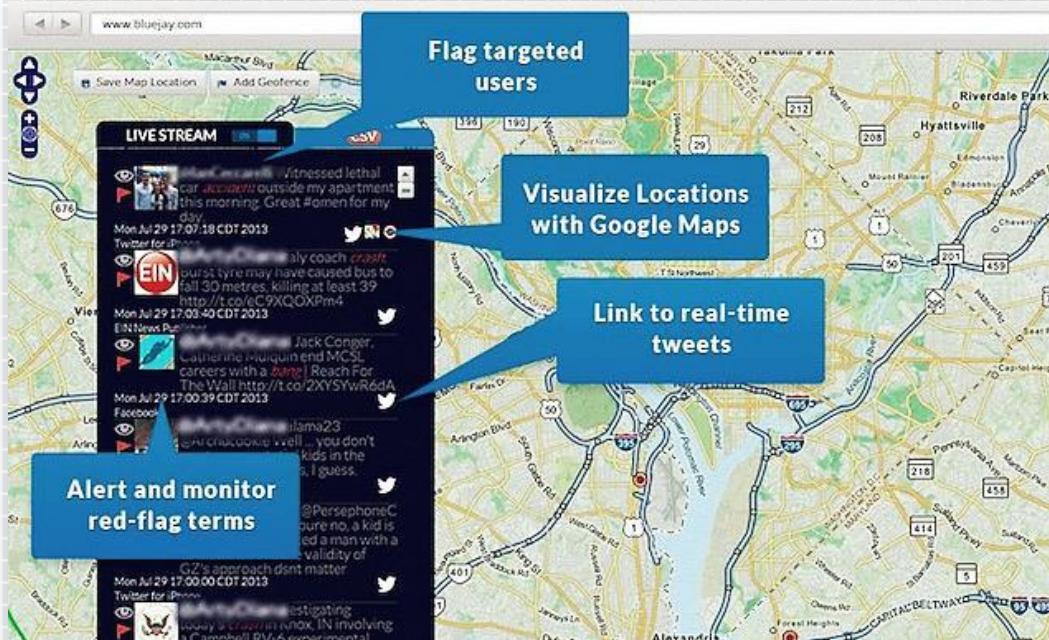
HUNT AND GATHER



Found **seed set** in under 3.6 hours on Friendster social network consisting of 5.6m nodes and 28m edges

2013 US Army ‘Viral’ Marketing Tipping Model

HUNT AND GATHER



2013 US Police (From BlueToad to BlueJay)

<http://arstechnica.com/tech-policy/2013/09/how-the-cops-watch-your-tweets-in-real-time/>



“We're not telling you how to do police work. We're just telling you the **best locations for you to be at** any given time of the day.”

-- Deputy Chief on Predictive Policing



The Humanitarian Data Exchange

WHERE YOUR DATA COMES TO LIFE

FIND DATA



e.g. Colombia or Ebola

Choose Location

1,580

DATASETS

243

LOCATIONS

207

SOURCES

2015 PHONE METADATA AS PROXY

Food Security and Poverty Indicator

FOOD ITEM (VARIABLE)	CORRELATION RANGE
Vitamin-rich vegetables (carrot, orange, sweet potato), rice, wheat, bread, sugar, meat	[0.7–0.8]
Eggs, oil, milk, butter, organ meat	[0.5–0.6]
Sorghum, ground nuts, seeds, fish, fruits, cooking banana, green leafy vegetables, beans, peas, maize, white roots, tubers, pumpkin, squash, cassava	[0.0–0.4]
White sweet potato	-0.4



Privacy Caveat: “Food items cultivated for personal consumption like cassava and beans show no relation with airtime purchases.”

Before: 3-5 Year Intelligence Delay
 After: Real-Time Human Safety Trend Warning

2013 STORY-DRIVEN VIDEO SUMMARY

1 Serving Ice Cream

Our method



2 Weighing Ice Cream

Uniform sampling



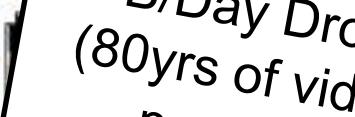
3 Watching TV in Ice Cream Store

Shortest path



4 Driving Home

Object-driven



Real-Time Story Telling

2012
6PB/Day Drones
(80yrs of video per day)

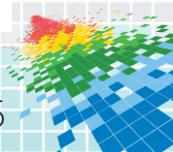
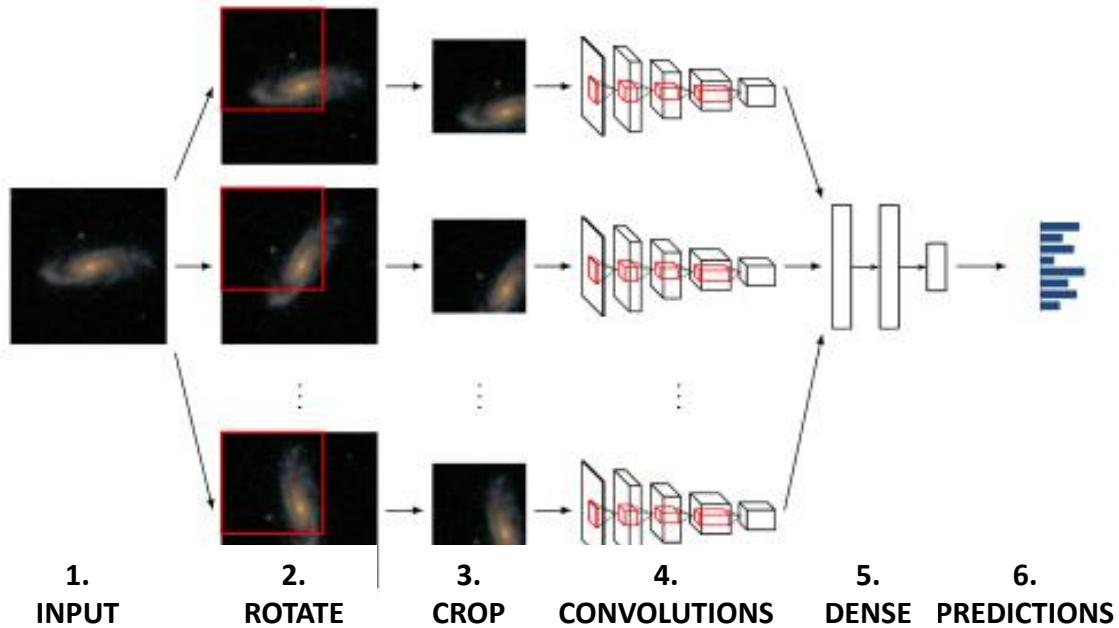
2015 DEEP CONVOLUTIONAL NEURAL NETS

Galaxy Morphology Prediction

1. Increased Compute Power Capability

2. *Human-Enhanced* Training Set Quality

Dieleman, Willett & Dambre <http://arxiv.org/abs/1503.07077>



**HUMAN
ENHANCED?**

2013 UNICEF DIGITAL MAPS

flickr

Sign Up

Explore

Create

Upload



Sign In



Voices of Youth Urgency Rank Testers

UNICEF GIS

PRIORITIZE ISSUES AND REDUCE DISASTER RISK

View all albums

Photos

15
Photos

397
Views



Deforestation in Kakolo township of Kitwe, Zambia - Brighton K.

EMC²

<http://unicefstories.org/2013/11/14/voices-of-youth-takes-digital-mapping-to-the-next-level/>
<https://www.flickr.com/photos/unicefgis/sets/72157636166333326/>

RSA Conference 2015



HUMAN “INFO SHARING” FOR SURVIVAL

430 BCE Athens (20% die)

Thucydides : “those who survive are immune”

910 CE Baghdad

Rhazes: “Transmitted from person-to-person. Theory of acquired immunity”

1022 CE China

Buddhist Monk variolation

1717 Istanbul (60% infected, 20% die)

Lady Mary, a British secular writer, observes “*ancient Turk practice*” of variolation – inoculates 6yr-old son with smallpox matter

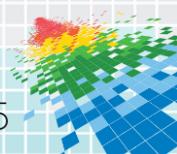
1721 English royal family tests variolation on prisoners and orphans

1765 London Medical Society paper “Cow pox and ability to prevent smallpox”

1774 English famer Jesty tests cow pox vaccination on wife and sons

1796 English Dr. Jenner observes rural customs, tests on gardener’s son

1840 Britain bans variolation



DR. JENNER'S VACCINE STORY (PROOF)

A = Smallpox Infection (30% chance of death)

B = Variolation (1-2% chance of death)

C = Cow pox Infection

V = Vaccine (Immunity to A)

“Vaca” means
cow in Latin



- ◆ Theory: C = V (Given C better than B, B better than A)
- ◆ Test: C = V if B after C has no signs of A
- ◆ Result: V = (or better than) C

INTELLIGENCE TRADE-OFF

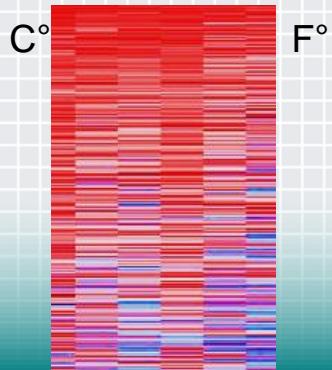
BINARY



RANKED



MEANING



0 POINT



EXACT



FREEDOM
(ERRORS)

INTELLIGENCE

LESSONS FROM IMPERFECT INTELLIGENCE...

(THREATS)



CONTEXT: OVERTRAINING



...they saw him coming on their radar screens but thought he was probably a flock of geese.



CONTEXT: NUANCE

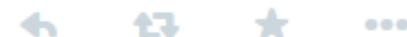




Matt Roller
@rolldiggity

New Parent Idea:

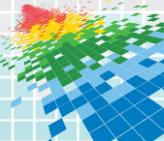
1. Take pictures of you pulling baby out of spacecraft in forest.
2. Hide pictures in attic for kid to find when he's 10.



RETWEETS FAVORITES

13,652 7,580

10:09 AM - 9 Apr 2013





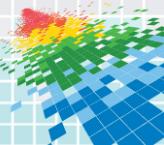
2015 April Fools Joke

400K shares traded in 5 min ... heaviest minute of trading volume since Feb 12th first minute of trading



I smile when
I'm faced with
a very absurd
situation,
when I can't
prevent my
own emotions.

– Benigno Aquino
Philippines President



CONTEXT: PROBLEM SOLVING



US officials 'aren't competent to accurately interpret tattoos.'

– J. Boerman, Expert Witness on Gangs



The Biometric Center of Excellence is seeking...any current databases containing tattoo/symbol images, their possible meanings... (Aug 2012 FBI RFI)



Age Range		Age 26 - 27
Gender		Female
Ethnicity Based on Surname		American
Education		Completed College
College Graduate		True
Marital Status		Married
Presence of Children		Children Present
Number of Children		2 Children
Children's Age		14 Years old, 17 Years old
Children's Gender by Age		Unknown Gender 11 - 15, Unknown Gender 16 - 17

“...a 26-year-old mother of two teenagers...is just about biologically impossible”



Wrong

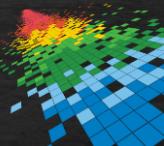
CONTEXT: EMPATHY



REAL EPISTOMOLOGICAL DANGER



“...technology built for counter-terrorism
used against human language, because
that’s where intentions are shown...
...get people before they act”



Weather

Partly cloudy tonight and Friday.
Cloudy Friday.

NIGHT EDITION

Only Cincinnati Paper Having Its Own
Independent & Complete Service.

10 Pages

No Cincinnati paper has been
able to give news of the Earth-
quake as quickly and accurately
as The Post.

VOL. 51. NO. 92.

CINCINNATI, THURSDAY, APRIL 19, 1906.

PRICE ONE CENT.

The Scripps-McRae Press As-
sociation has had the first and
best news of the Earthquake
from the outset.

THE CINCINNATI POST

3000 DEAD; \$300,000,000 LOST

SAN FRANCISCO IS OBLITERATED

THIS SHOWS THE EXTENT OF THE DEVASTATION IN SAN FRANCISCO



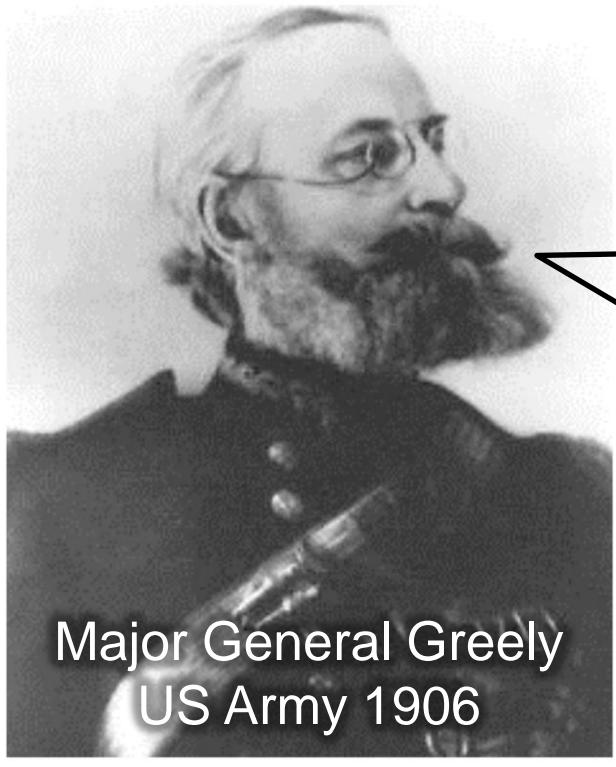
PHOTOGRAPHED FROM NOB HILL, RESIDENCE SECTION, NOW ON FIRE. ARTIST'S WORK SHOWS BIG BUILDINGS AND LANDMARKS AND THEIR FATE.

Thursday Sees Survivors Facing
Death From Hunger and Thirst.

**EARTHQUAKE IN A LATEST BULLETINS
NUTSHELL**

Frisco Transformed From a City of
Joy to a Place of Death

HUNDREDS DIED



Balderdash!

I say official SF
number dead is

498

“Come on, let us shoot...”

US Apache helicopter
July 12, 2007



Killed:
Reuters photographer Namir Noor-Eldeen and driver Saeed Chmagh...

<http://www.reuters.com/article/2010/04/06/us-iraq-usa-journalists-idUSTRE6344FW20100406>

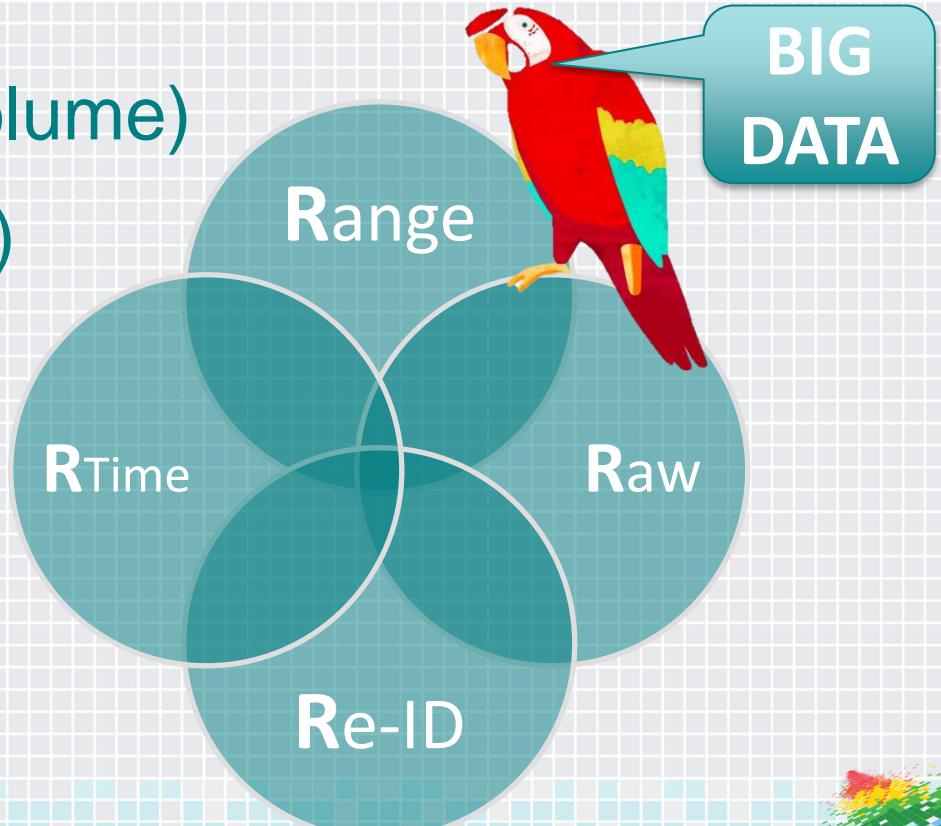
GIVEN CONTEXT AND THREAT... WHAT CAPABILITIES EXIST?



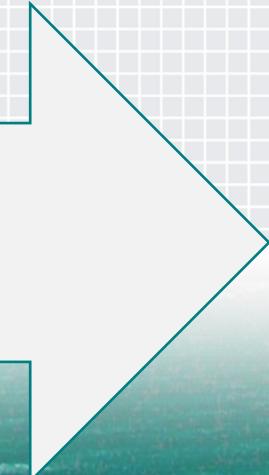
#RSAC

THE RRRRs of “MASSIVE SCALE”

1. Range of Sources (Volume)
2. Raw Formats (Variety)
3. Real-Time (Velocity)
4. Re-ID (Vulnerability)



MASSIVE SCALE = “DATA LAKE”



DATA CONTINUITY METHODOLOGY

Easily Identified

Identify Sick ASAP

Routine Treatment

Store Adequate Records

Minimum Judgment

Evaluate Daily Sick

Adapt Until Noted Improvement



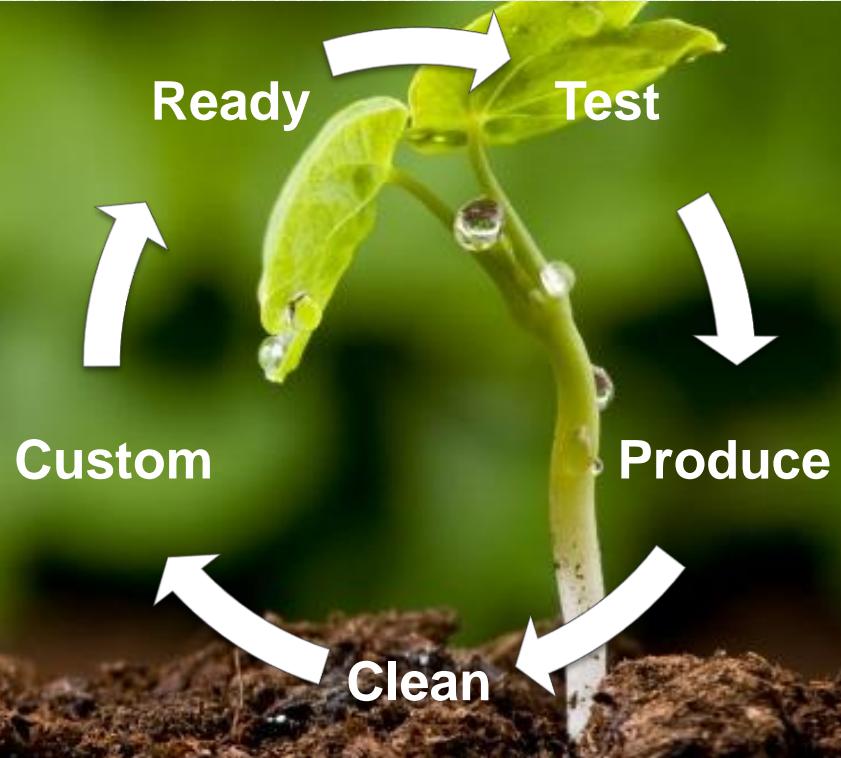
ERM IoT SURVIVAL (Ringworm Neuron Circuit)

If large food decrease, release dopamine

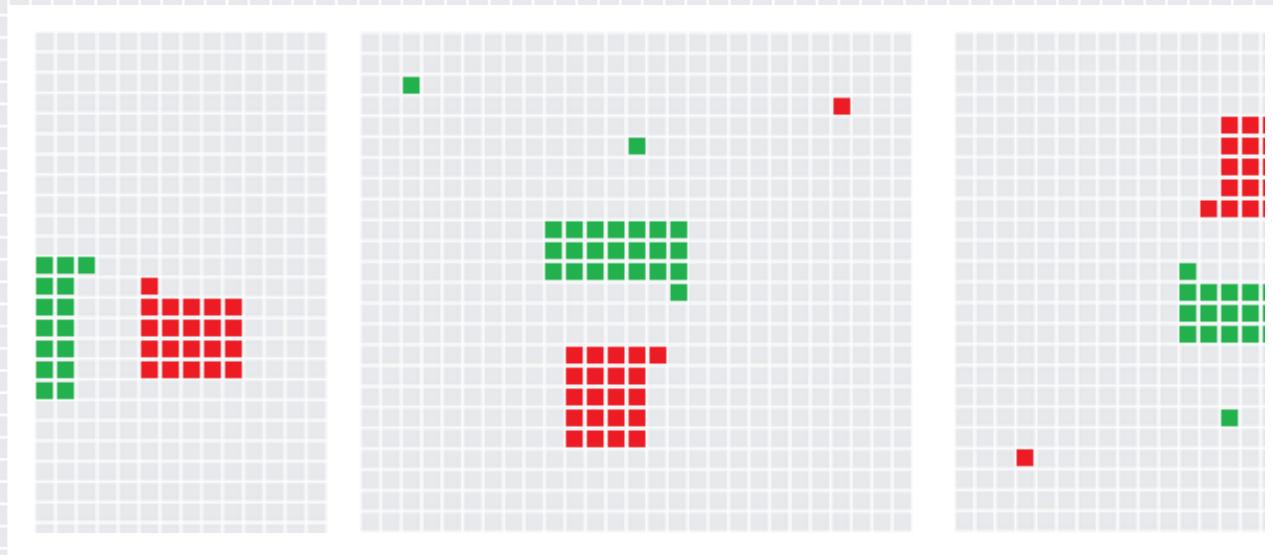
If dopamine released, take more risks (expand search for food)

ISEA SYSTEMS SURVIVAL PLATFORM

“Service Compost”



ISEA SYSTEMS SURVIVAL ACTION

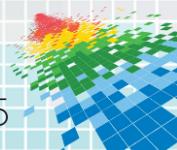


INTELLIGENCE EVOLUTION



PEOPLE, DATA and TRANSACTIONS

	Artisanal	Industrial	Environmental
Focus	Verify Identities	Change Perimeters	Examine All Behavior
Capability	Technology	Procedures	Value
Engagement	Single Well-Defined Events	Closely Related Events	Isolated Events (Correlated)
Output and Outcome	Isolated; Disengaged	Engaged In; Disengaged Out	Engaged In & Out
	Areas of Innovation, Automation, Fixes	Services, Solutions and Dashboards	Risk Innovation; New Types of Business



COMPUTE POWER REVOLUTION

1996

Artisanal

“hardware cost for **24-node BEOWULF** cluster \$57K —compared to commercial supercomputers between \$10m and \$30m”



2006

Industrial

“**30** men's
heft of grasp
in the gripe of
his hand”

SMARTER, BETTER, FASTER

◆ Find: Improve Identification

1. Human Enhancement
2. Compute Power

◆ Survive: Data Continuity Method and Technology

2016?

Environmental

Examine All Behavior

Value

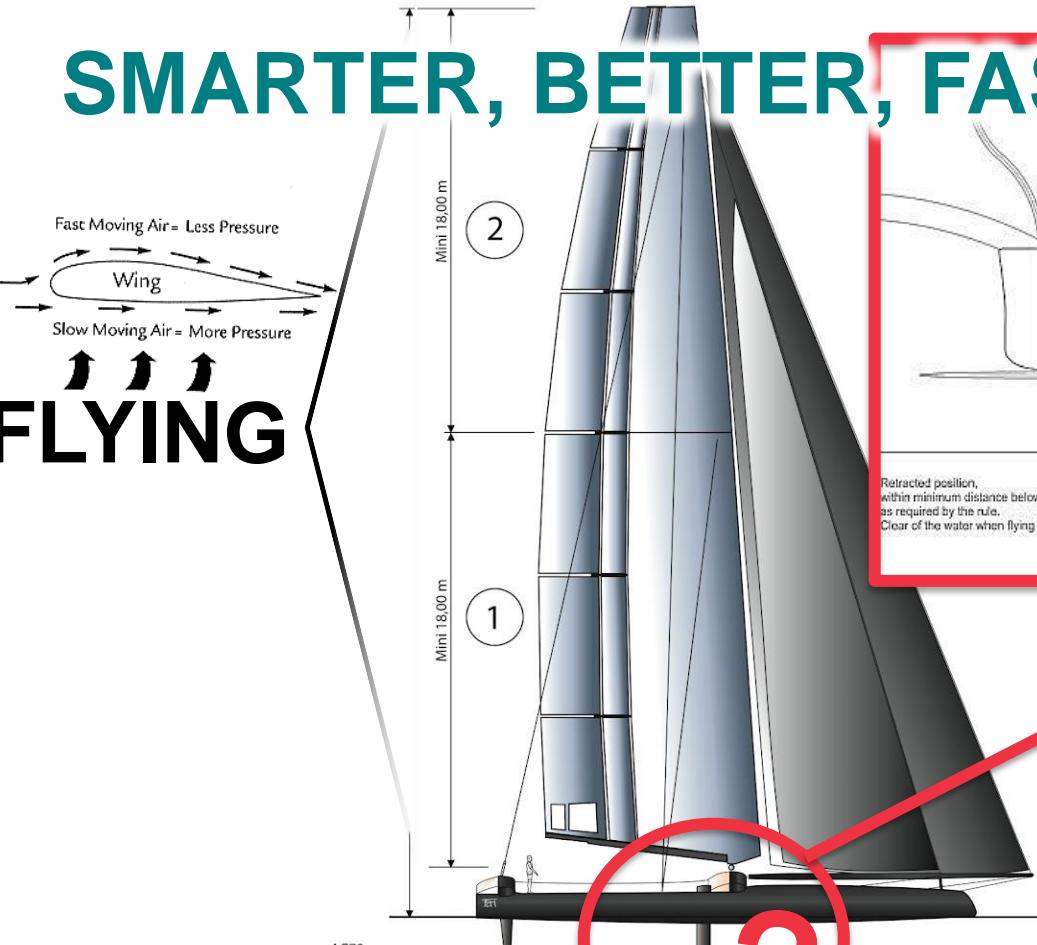
Isolated Events
(Correlated)

Engaged In & Out

Risk Innovation; New
Types of Business

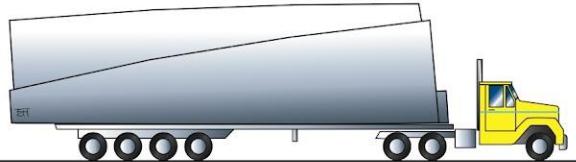
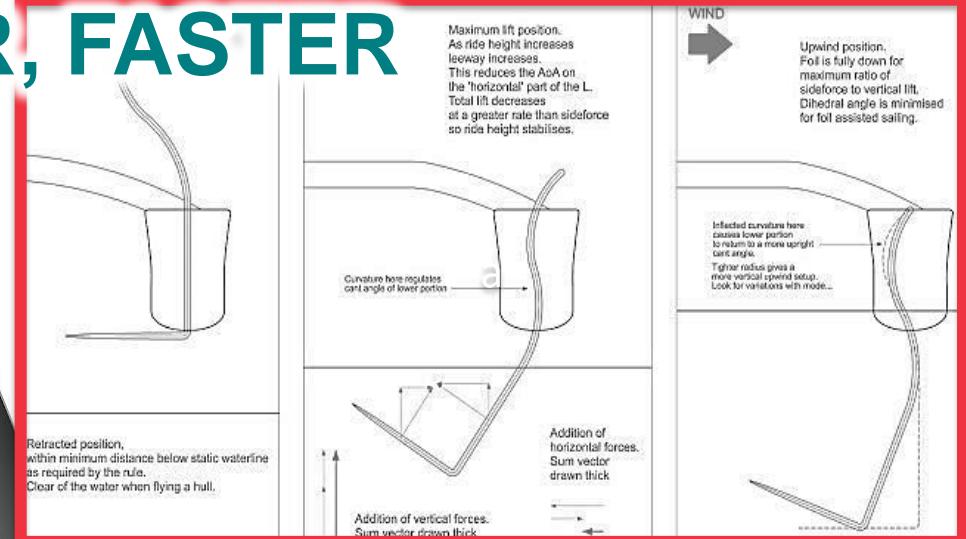
SMARTER, BETTER, FASTER

Fast Moving Air = Less Pressure
 Wing
 Slow Moving Air = More Pressure
FLYING



AC72
 Francois Chevalier YD
 30/05/2012

“Little Herbie” (Stability Augmentation Systems)



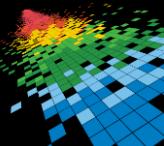
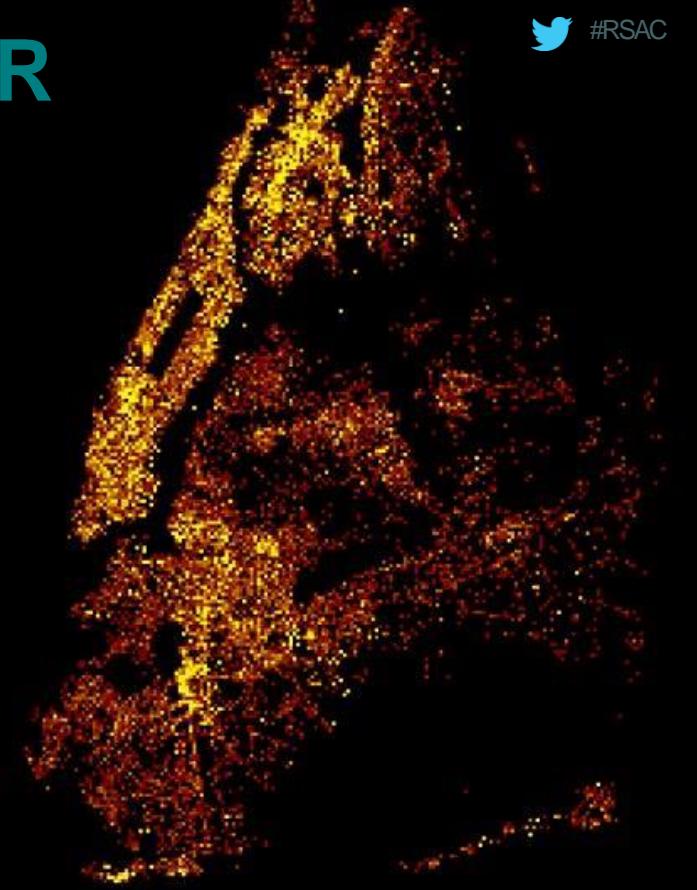
SMARTER, BETTER, FASTER

no bubbles

6 mph

bubbles
20+ mph

SMARTER, BETTER, FASTER

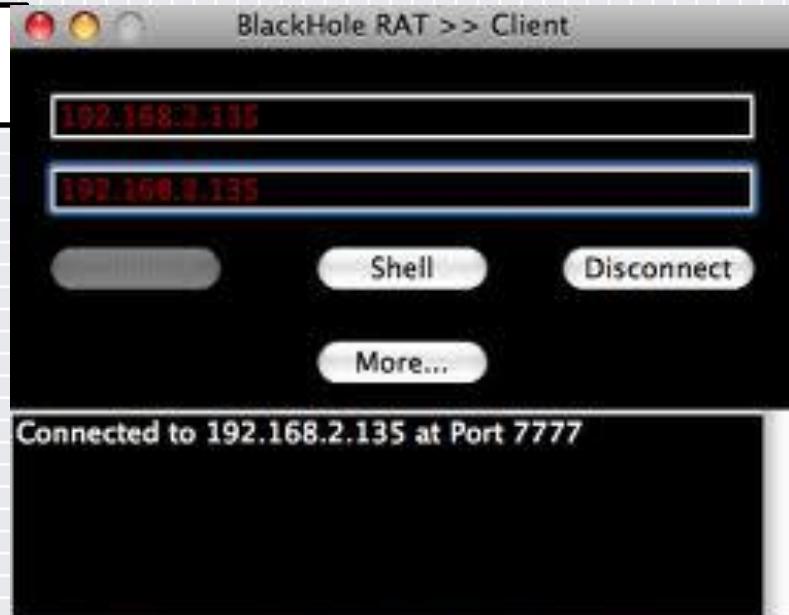


SMARTER, BETTER, FASTER



Average 2 Burritos/wk	145lbs
Campaign Donation	Running RAT (fingerprint)
2 Children	Flight to Vegas
Called Doctor x5	231 Yrs Old
Patched Chrome	Streaming Blade Runner
Unpatched Android	

"I don't like that man...have to get to know him better."



SMARTER, BETTER, FASTER

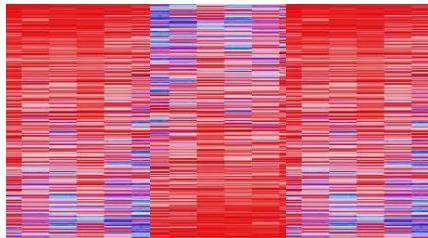
Context	IT Capability Required	Average
E-mail	Address	2
Social Media	Username	3
Telecom	Phone Number	2
Finance	Account Number	4
Health	Account Number	3
Government	Identity Number	3
E-Commerce	Username	5
Work	Username and Badge	2

24

TYPICAL IDENTITY COUNT PER USER

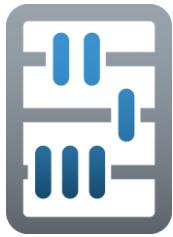
SMARTER, BETTER, FASTER

Harvesting
and Indexing



“Big Data Haiku”

Ingest, Store, Analyze



“Sqoop”

EMC²

Surface
(Web Apps and Services)



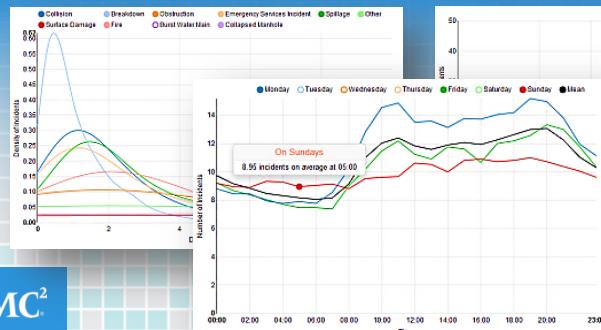
“Sqoop”

EXAMPLE: LONDON TRAFFIC ANALYSIS

Bakerloo	Suspended
Central	Suspended
Circle	Suspended
District	Suspended
East London	Suspended
Hammersmith & City	Suspended
Jubilee	Severe delays
Metropolitan	Suspended
Northern	Good service
Piccadilly	Part suspended
Victoria	Suspended
Waterloo & City	Suspended

Traffic Disruption
Data Feed

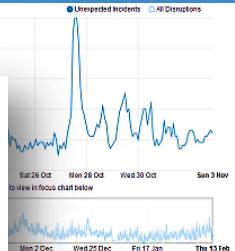
Surface



EMC²



d3.js & NVD3
Interactive SVG Figures



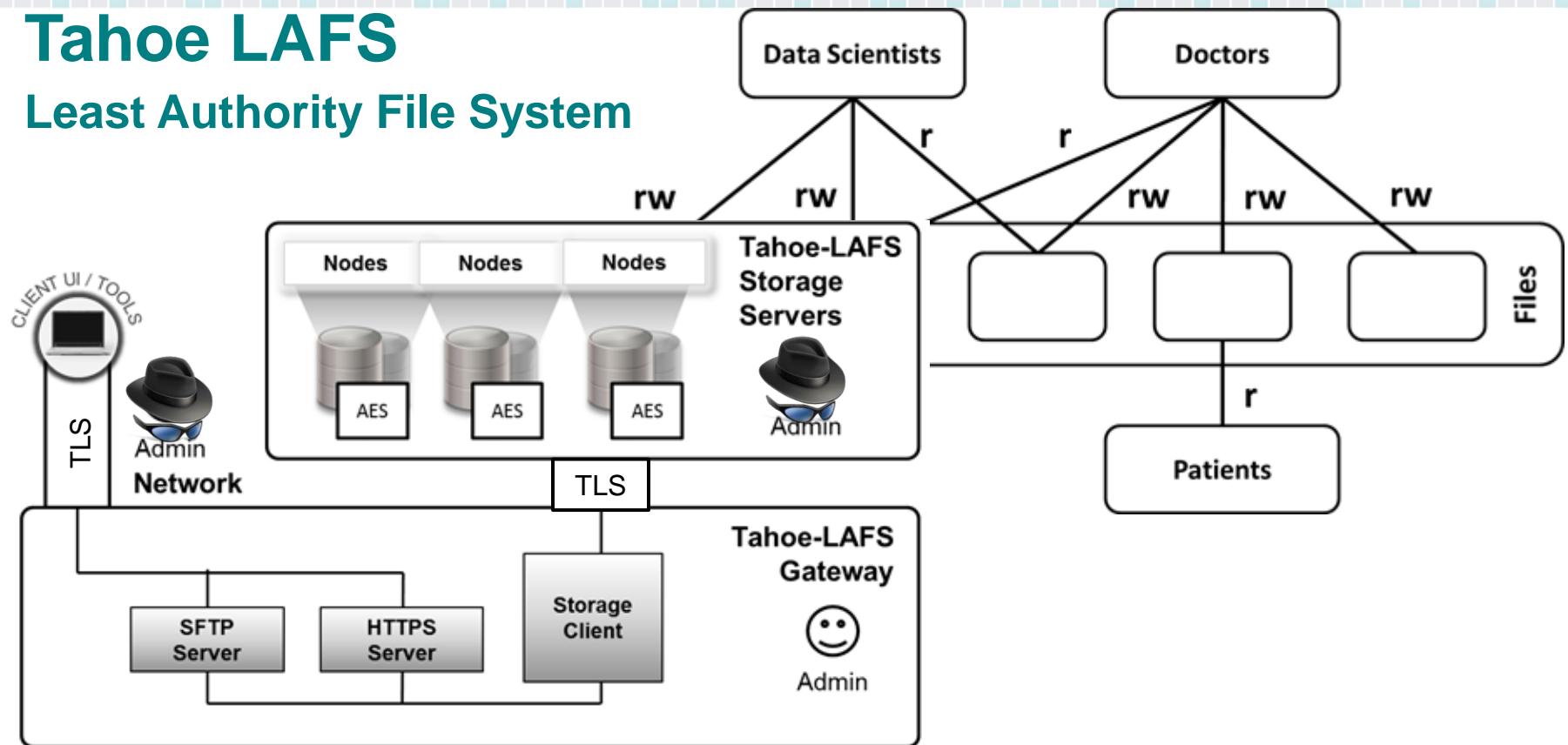
Analyze

scikit-learn MADlib python

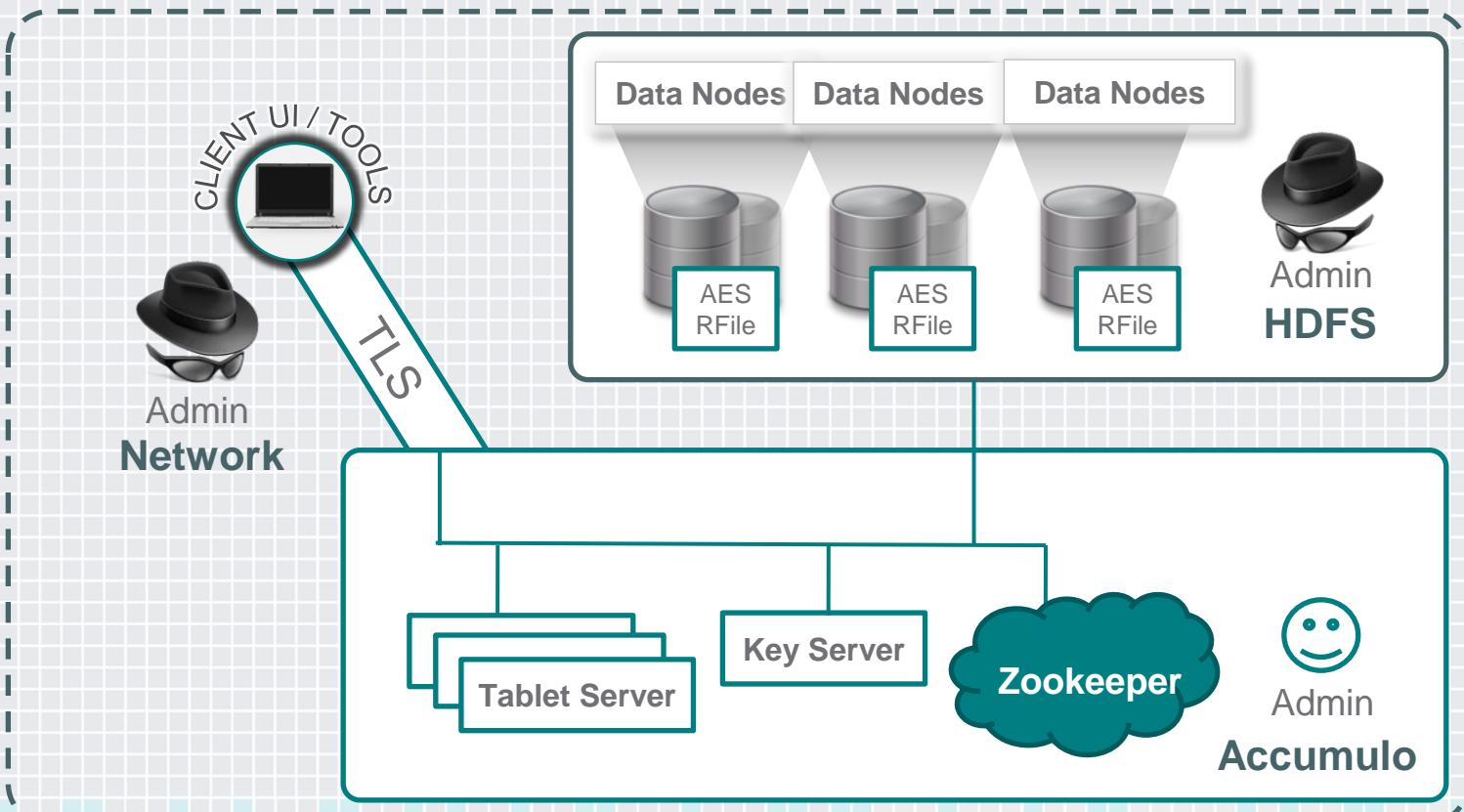


ADVANCED PERSISTENT Tahoe LAFS

Least Authority File System

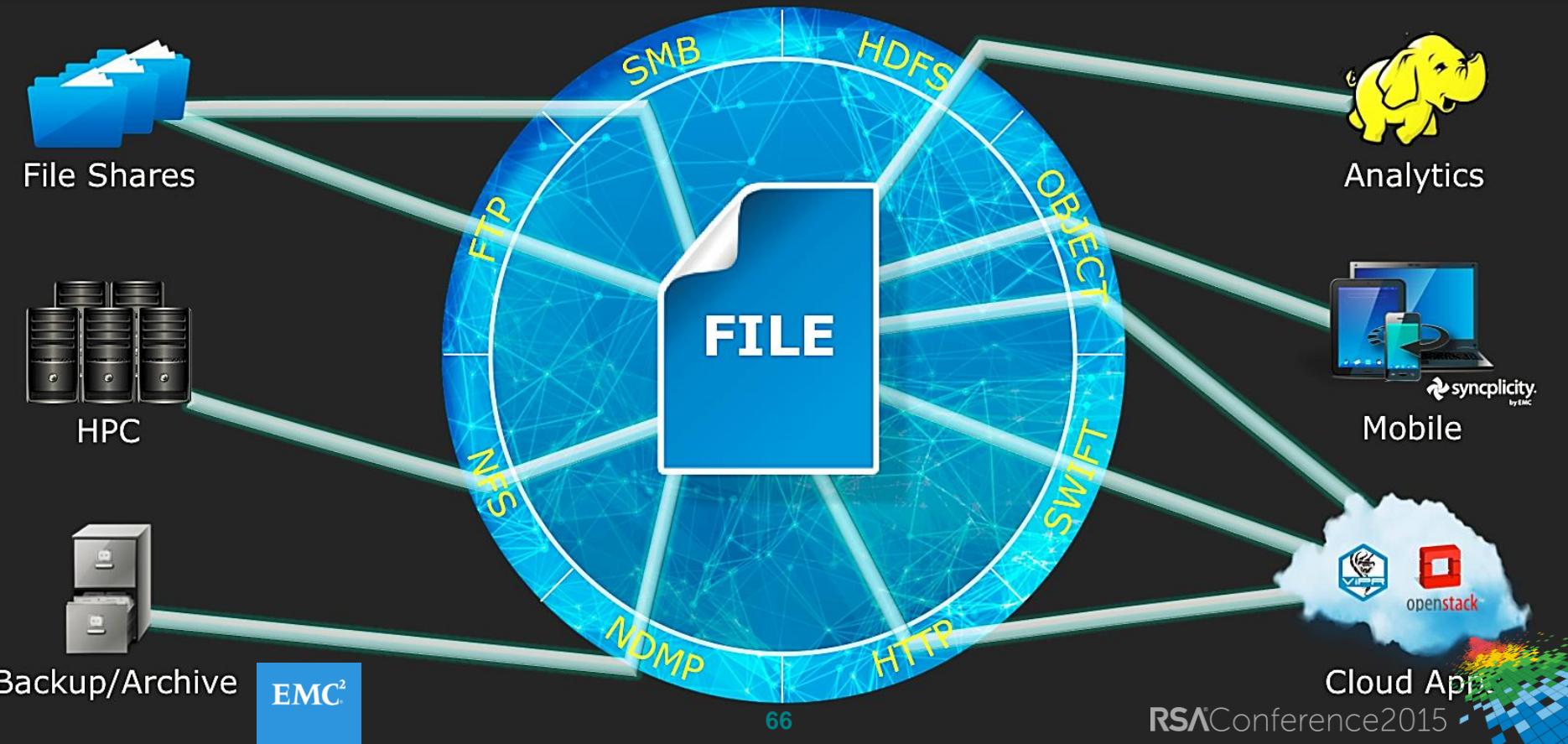


ADVANCED PERSISTENT ACCUMULO



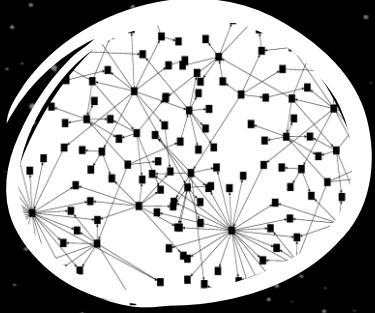
TEXTUAL DISTRIBUTED DATA CONTROL

#RSAC





+



+



CONTEXT

Persuasion, Empathy,
Problem Solving,
Negotiation, Physical Skill

THREAT

Familiarity, Change,
Persistence

CAPABILITY

Volume, Velocity,
Variety,
Vulnerability

RSA® Conference 2015

San Francisco | April 20-24 | Moscone Center

SESSION ID: DSP-W01

MORE APT THAN YOU THINK: DATA PROTECTION AT MASSIVE SCALE

Davi Ottenheimer

Senior Director of Trust, EMC
@daviottenheimer

Wednesday, Apr 22nd 8:00 AM | West 3006

CHANGE

Challenge today's security thinking

