

RAeS Flight Simulation Group Conference
12/13 June 2018 - London

It's Big Data but Where Is It?

**Towards an Enterprise Approach to Training
Data and Analytics**

Andy Fawkes



A Metaphor for Your Enterprise Training Data



Unconnected Small Data?

or



Big Data?

Overview

- **What is Big Data?**
- **Big Hype?**
- **Exploiting Big Data in Training**
- **Challenges**
- **Way Forward**

Presentation has a Defence Emphasis



What is Big Data?



What is Big Data?

“data sets that are so voluminous and complex that traditional data-processing application software are inadequate to deal with them”

Big Data - 4 Main Features



Volume



Velocity



Variety

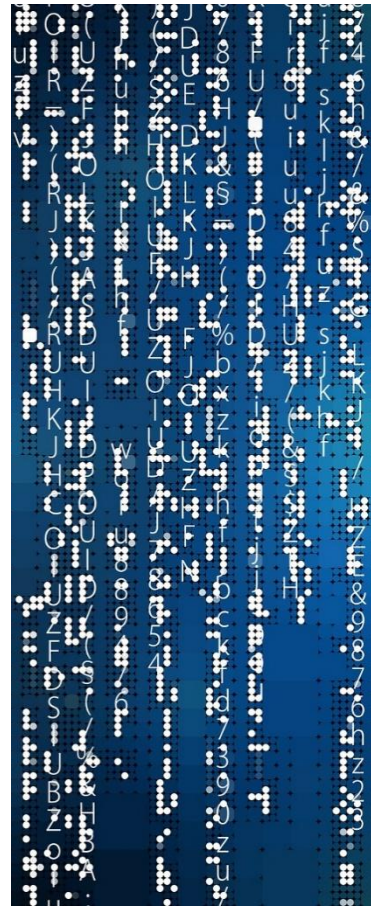


Big Data - Ultimately it's...

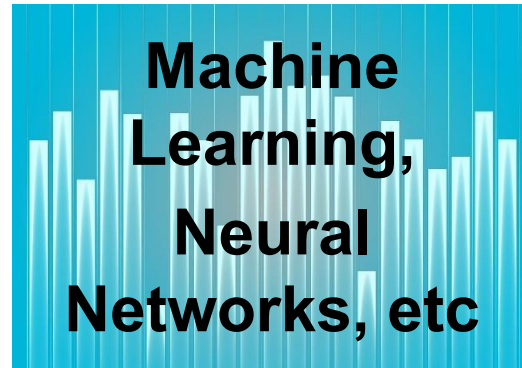


Creating Added Value at the Right Time

Exploiting Big Data



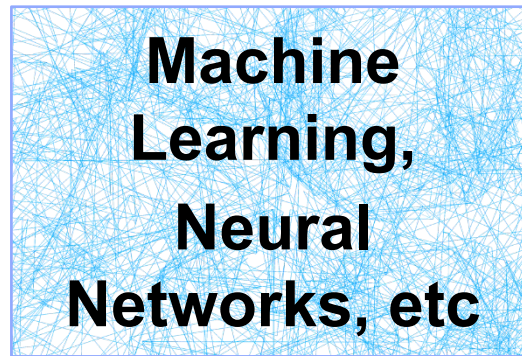
Big Data



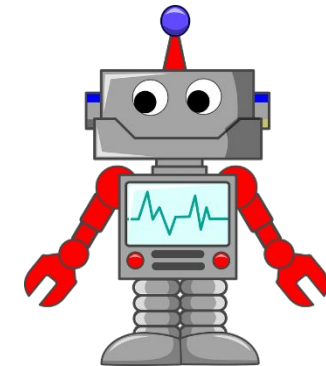
Analysis



Human Decisions



“Training”



Automation

Why Big Data in Analytics?

Prescriptive Analytics

- seeks to determine the best course of action

Predictive Analytics

- identify past patterns to predict the future

Diagnostic Analytics

- used to determine why something happened

Descriptive Analytics

- what is happening now based on incoming data

Big Hype?

A Little History First

1944 - Fremont Rider Research

Predicted American libraries would double their capacity every 16 years

1975 - Japan Govt. Study

One-way communication transitioning to two-way

2000 - Peter Lyman/Hal R. Varian

Estimated the world generated about 250MB per human per year

2009 - Roger E. Bohn Study

Estimated Americans consumed information for almost 12 hours per day

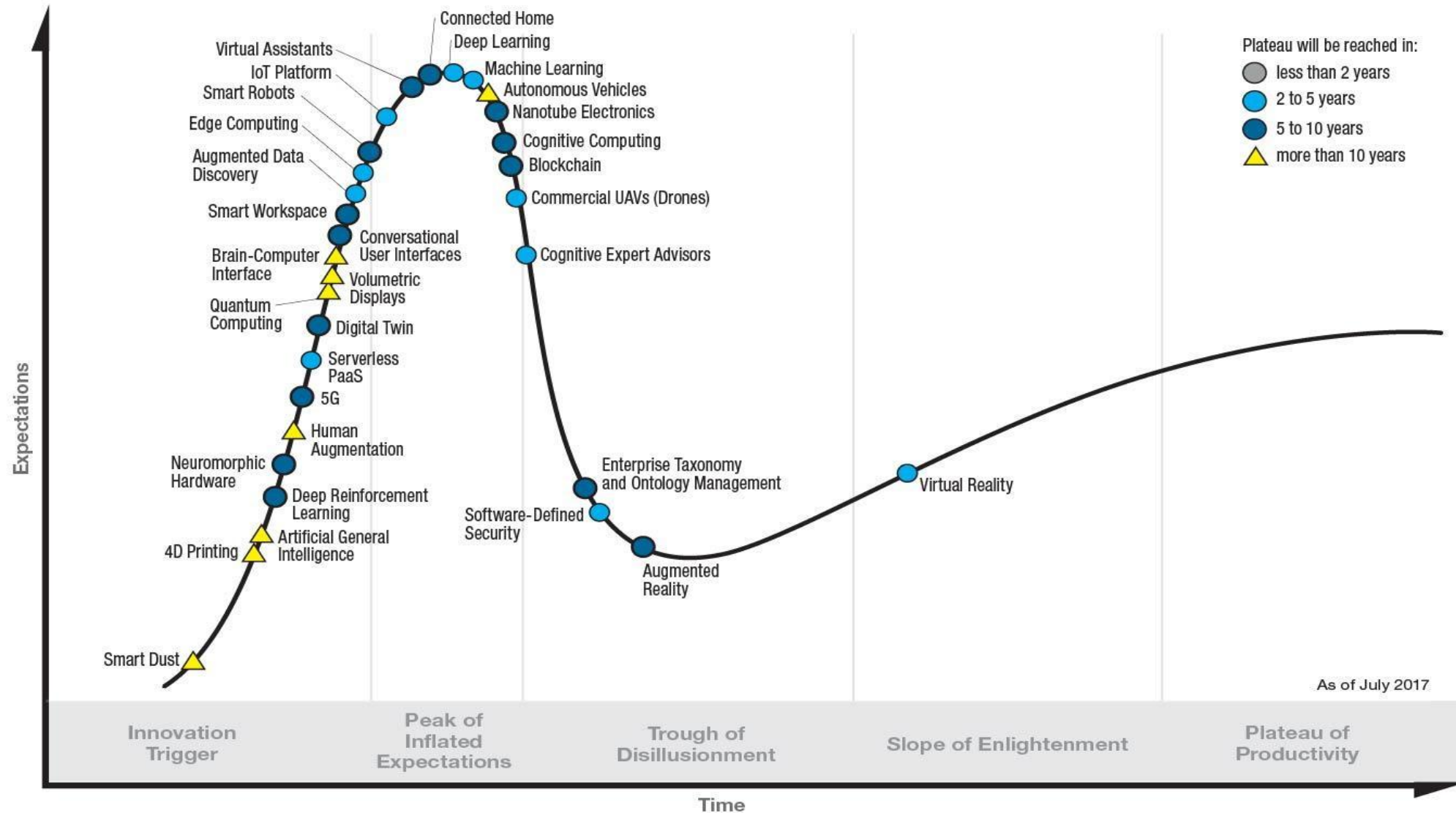
2018 – Hootsuite Study

Estimates that there are 4 billion people around the world using the internet

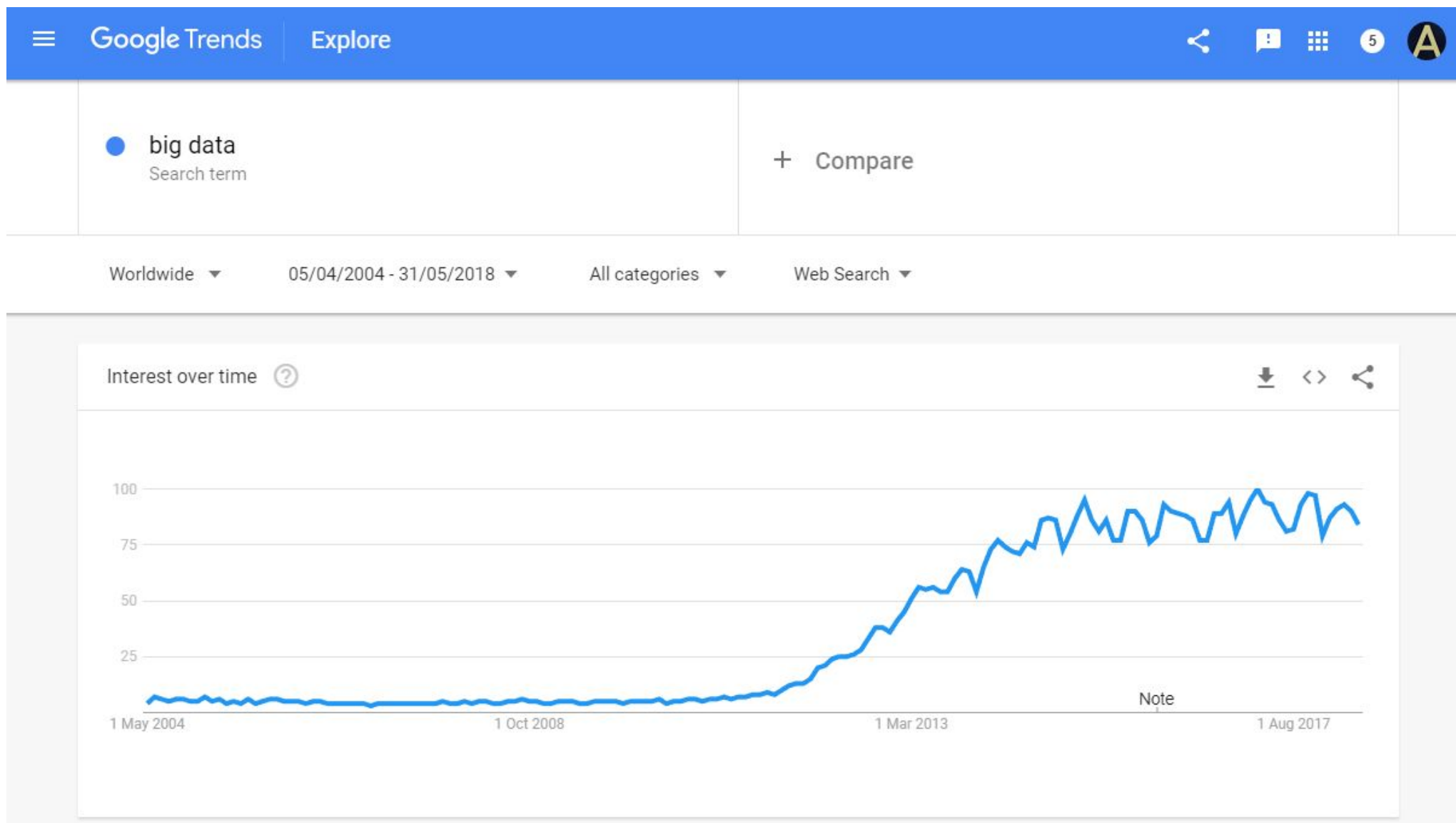
2020 – IDC Research

Predicts 54TB new information generated for every human per year

Gartner Hype Cycle - Emerging Technologies 2017 - Big Data?



“Big Data” as a Search Term



Big Data in the Economy



[Departments](#) [Worldwide](#) [How government works](#) [Get involved](#)
[Policies](#) [Publications](#) [Consultations](#) [Statistics](#) [Announcements](#)

[Home](#)
[Press release](#)
New national innovation centre to put UK at forefront of big data

A new £30m National Innovation Centre for Data (NICD) aims to see the next Google or Facebook started in the UK and help the country capitalise on a potential £40bn a year boost to the economy.

Published 14 September 2017
From: [Department for Digital, Culture, Media & Sport](#) and [The Rt Hon Matt Hancock MP](#)



Related content

[Policy](#)
[Business enterprise](#)
[Cyber security](#)

“We’re determined to unlock the huge potential of big data which could add billions of pounds to our economy - from powering price comparison sites to improving the flow of transport around cities.”

Minister of State for Digital
Matt Hancock MP
14 Sep 2017



Defence Information Strategy

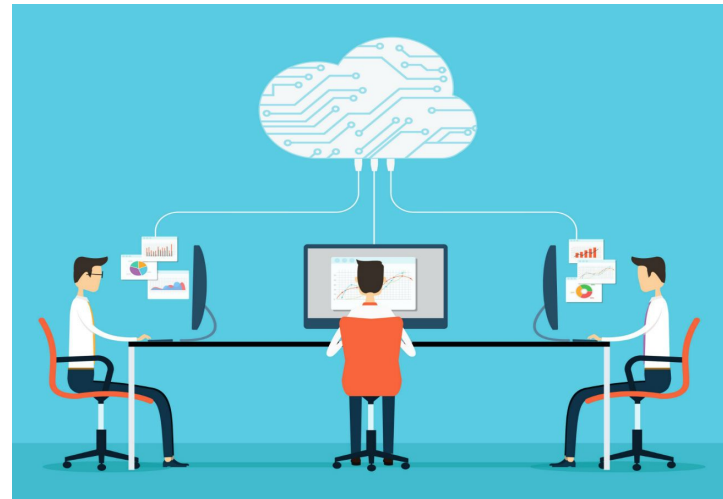
28 Dec 17

Defence needs to exploit the full potential of the data it holds through active data management, better digital processes, and more timely and cost effective analysis to derive maximum value to support evidence-based decision making.

Defence will address the 'big data' challenge by developing the enterprise technologies, skills and strategic partnerships, to allow data analytics to be exploited by the business to make better decisions.

Exploiting Big Data in Training

Training Data Generators



Different Human Related Systems

- Personnel Records
- Education Records
- Training Records
- Training System Records

What if they could be brought together?

Big Data in Training

Prescriptive Analytics

- eg. cost effective balance of investment in training

Predictive Analytics

- eg. linking recruitment data to pilot success rates

Diagnostic Analytics

- eg. why are failure or success rates so high?

Descriptive Analytics

- eg. linking readiness to humans and aircraft

Other Benefits

- **Enterprise**
 - **Benefits Increase over Time as Data Builds**
- **Personnel**
 - **Complete Service Record, Recruitment to Post-Military Career**

Challenges

Air Training Capabilities

A400M

Apache

ASDOT

AUV

C-130

C17

DCS&S

DOTC(A)

DSALT2

E-3

F-35

iASTC

LTPA

MAA

MFTS

MSHTF

P-8

STARS

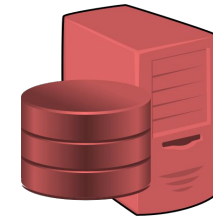
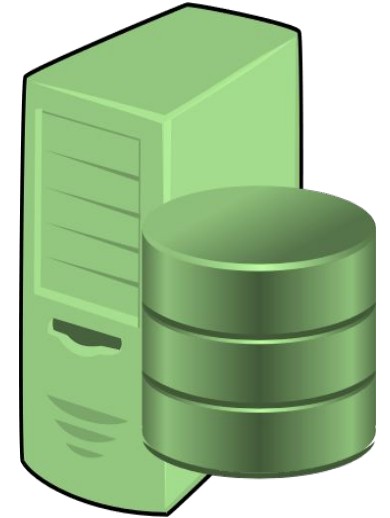
TFST

TyTan

WIST

Enterprise Training Data

Unconnected - Incompatible - Inaccessible



Temporal Data Management



Can
government
stop losing
its mind?

How network
thinking can help
us record, find and
use what we need

Gavin Starks
April 2018

nesta

- Can government remember?
- Is it condemned to repeat mistakes?
- Or does it remember too much and so see too many reasons why anything new is bound to fail?

Data Misjudgements



Google Street View

Data Dangers

Emergent Tech ► **Internet of Things**

You know that silly fear about Alexa recording everything and leaking it online? It just happened

US pair's private chat sent to coworker by AI bug

By Kieren McCarthy in San Francisco 24 May 2018 at 18:49 **221**  **SHARE** ▼



Updated It's time to break out your "Alexa, I Told You So" banners – because a Portland, Oregon, couple received a phone call from one of the husband's employees earlier this month, telling them she had just received a recording of them talking privately in their home.

“Amazon takes privacy very seriously.

We investigated what happened and determined this was an extremely rare occurrence.

We are taking steps to avoid this from happening in the future.”

Black Swan

Pre-1697



Post-1697



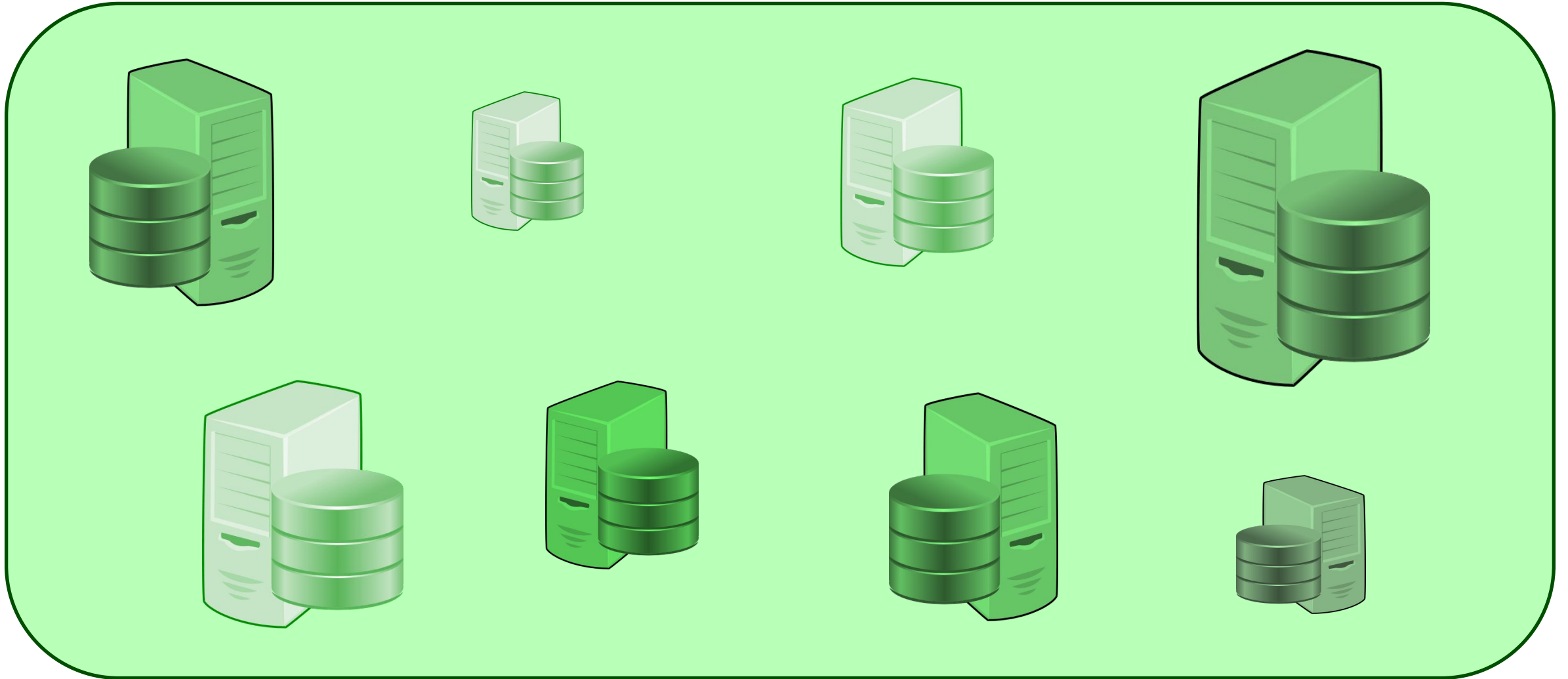
Black Swan Event

- The event is unpredictable (to the observer)
- The event has widespread ramifications
- After the event has occurred, people will think it was explainable and predictable

Way Forward

Enterprise Training Data End Goal?

Connected - Compatible - Accessible



Enterprise Training Data

- **Set Enterprise Training Data Requirements for all Projects**

and/or

- **Develop/Exploit Data Integrating Software**

Embrace Human Imagination

the surprises of the future are, by definition, not repeats of surprises in the past



THINKE

What is Your Approach to Enterprise Training Data?



Unconnected Small Data?

or



Big Data?

Questions?



THINKE