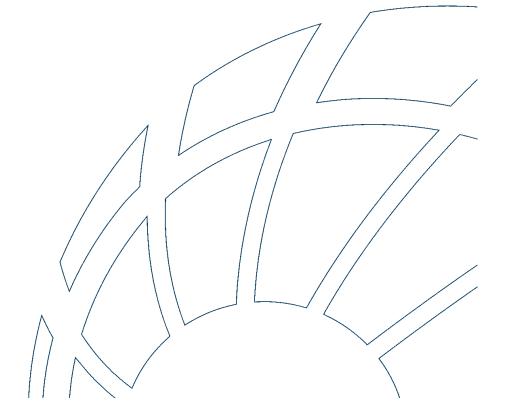


# Big Data

Beyond the hype

Benjamin Hodgson MBA PhD Senior research scientist BMT Group ESPO Conference 2-3 June 2016!

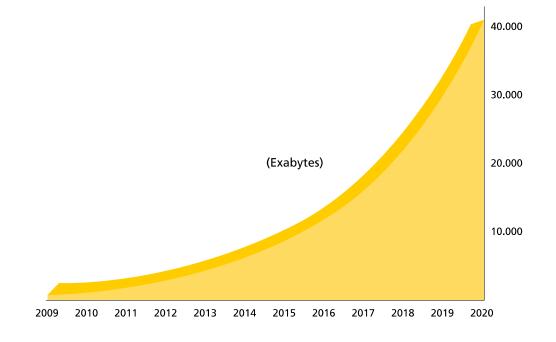


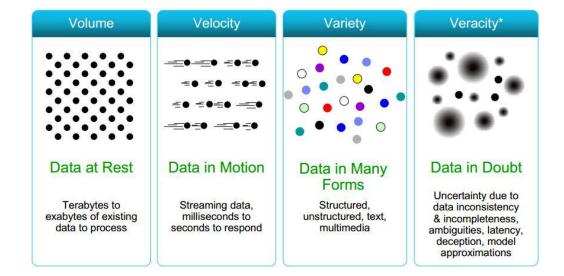


# What is big data?

Defining 'big data'

- 2001: Doug Laney at Gartner describes big data in terms of Volume, Velocity and Variety
- 2012: Gartner updates definitions to: "Big data is high is high-volume, high volume, high- velocity and high-variety information assets that demand cost effective, innovative forms of information processing for enhanced insight and decision making."







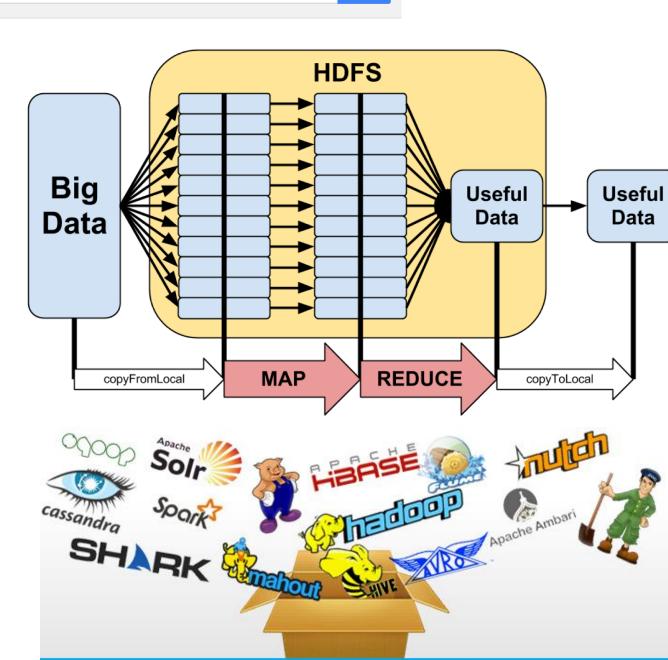




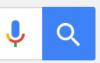
Press Enter to search.

The basic idea is to get some data Split it into chunks and process it in parallel.

- Clean it up
- Simple processing
- Statistics







# Your data is not that big

### Port of Rotterdam:

- 30,000 seagoing vessels per year.
- 110,000 inland vessels per year
- 12.2 million twenty feet equivalent unit per year

# Marine traffic (monitors AIS)

- 65/85k vessels visible at any time
- 20B positions archived.

# Your expertise is not ITC

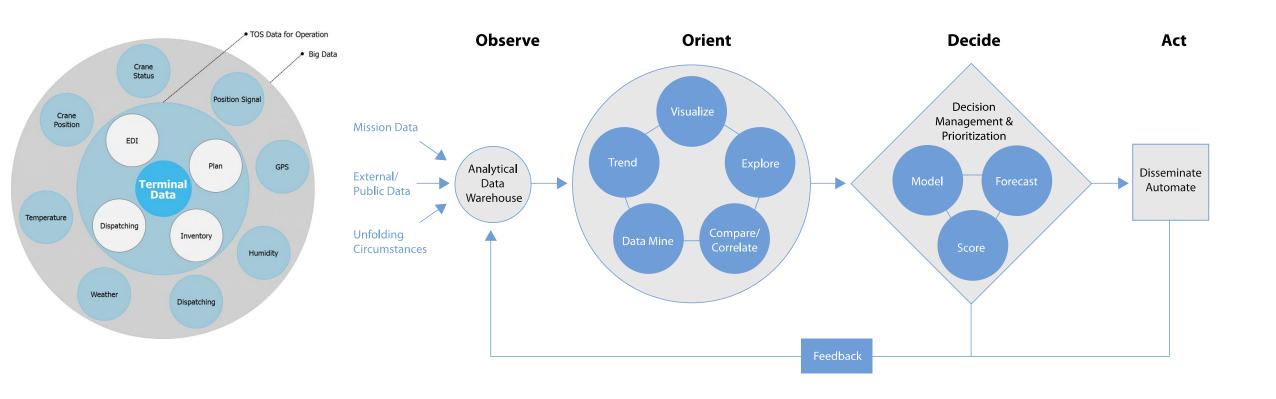
## Do you have skills in-house?

- Managing systems
- Data science

### Outsourcing carries risks.

- Critical systems.
- Strategic importance.
- Security

# It is just about data driven decisions.



Is there anything new here?

Just an enhanced TOS?

Just a normal database?

Just Operational Research rebranded?

### **Cynefin domains**

#### The domain of data sharing

- Partner
- Find mutual value.
- Exploit opportunities

#### **People not programs**

- Triage
- Visualisation

#### **COMPLEX**

Cause and effect are only coherent in retrospect and do not repeat

Pattern management

Perspective filters

Complex adaptive systems

Probe-Sense-Respond

#### **KNOWABLE**

Cause and effect separated over time and space

Analytical/Reductionist Scenario planning Systems thinking

Sense-Analyze-Respond

#### **CHAOS**

No cause and effect relationships perceivable

Stability-focused intervention

**Enactment tools** 

Crisis management Act-Sense-Respond

#### **KNOWN**

Cause and effect relations repeatable, perceivable and predictable

Legitimate best practice

Standard operating procedures

Process reengineering

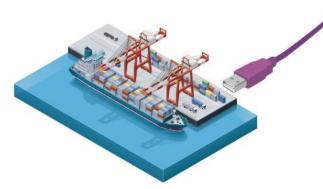
Sense-Categorize-Respond

#### The domain of modelling

- Modelling & prediction
- Use data to increase understanding.
- Causal relationships must be grounded in reality not just statistics.
- Small scale expert driven projects.

#### The domain of existing Automation.

- Extension of existing TOC/PCS
- Optimisation & automation.
  - Is the benefit worth the effort?



Big data is not magic it is just about making better decisions.

## Operational improvement.

- Have a clear idea of what you want to achieve
- Keep it simple you don't need a super computer.
- Test small scale before you invest.

#### Partner & Collaborate

- Your data is valuable to other stakeholders.
- Port and port community system is uniquely placed to integrate stakeholders.
- Be careful to respect legal and commercial confidentiality concerns