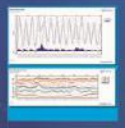
OceanWise

End-to-End Marine and Coastal Data Management and Decision Support

Intelligent Marine and Coastal Mapping Data

Data Policy, Strategy and Management Systems

Enterprise GIS and Productivity Tools



Environmental Data Sharing and Publishing

Capacity Building, Training and Mentoring

www.oceanwise.eu

Independent • Experienced • Innovative • International • Trusted

Marine Data & Information

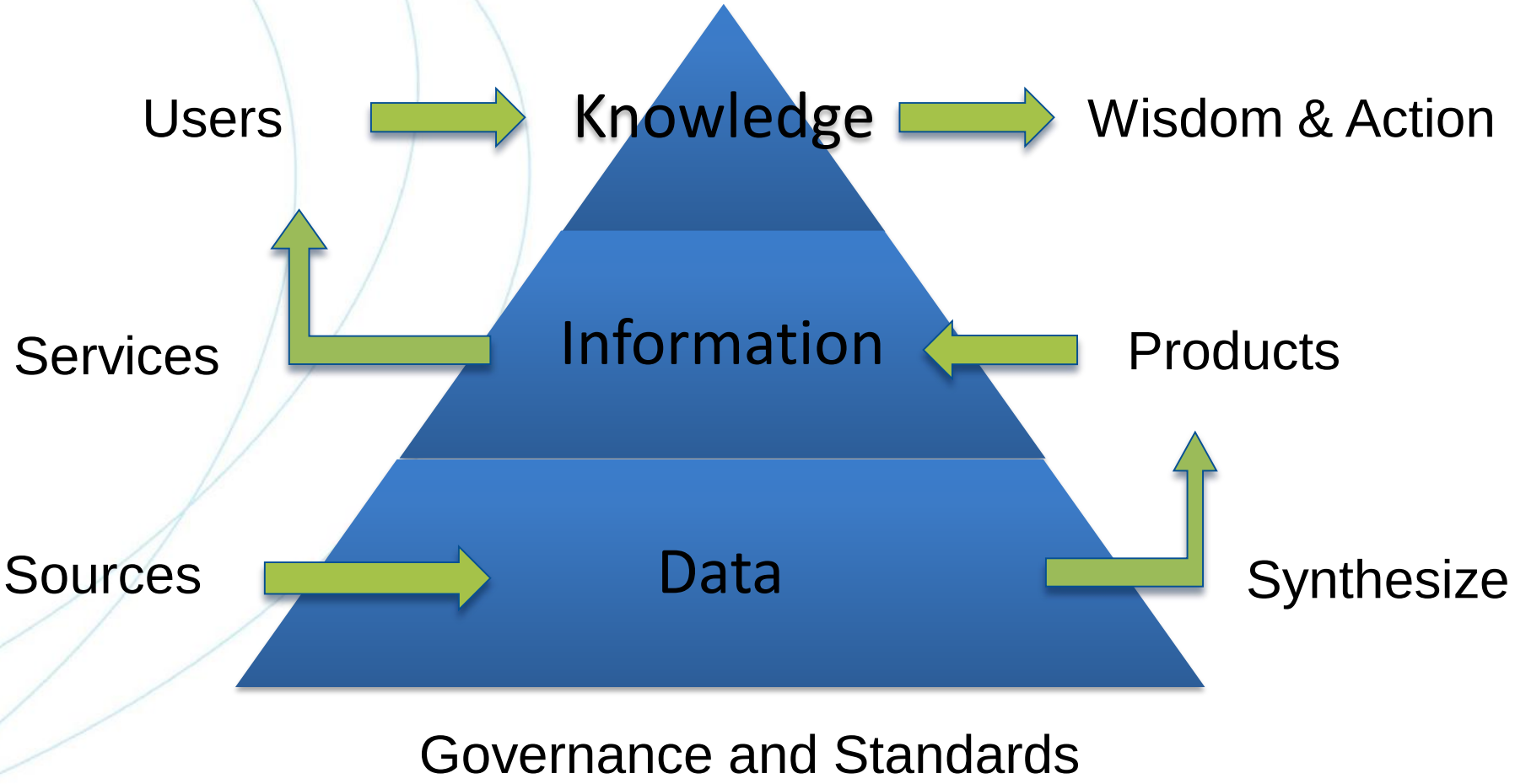
"Revolution and Evolution?"

John Pepper and Dave Morris
OceanWise Ltd

Key Messages

- **REVOLUTION**
  - ✓It's time to turn the way you think about data use upside down!
- **EVOLUTION**
  - ✓Knowledge Doubling and the Data Deluge
  - ✓Open Government and Transparency
  - ✓Open Data
  - ✓Linking Data
  - ✓Citing Data
  - ✓Changing the way we do things?

Data - Information - Knowledge - Wisdom (DIKW) Pyramid



REVOLUTION-What it means for Data

- The comfortable linear assumption that ***"the more data we have the stronger the base of the DIKW pyramid"*** leading to ever sounder Wisdom, Decisions and Actions is misleading us
- The rate limiting step isn't data, its not really analytics, it's not really anything to do with machines; **it's people**!
- We need to understand the provenance of all the processes, the assumptions, the arguments, the "if's and but's" and that means we need help from machines, from systems and from other people

REVOLUTION-What it means for Decisions

- We need to be much more aware of the ever increasing gap between our machines ability to acquire, process, model & display information and...
- our ability to assimilate, understand, theorize and use the resultant data products, (i.e. Information, Knowledge and Wisdom) to make decisions
- Inverting the DIKW pyramid would show our Knowledge and Wisdom to be at the base of a very unstable, non linear assemblage of stuff

OceanWise

HERE BE DRAGONS

HC SVNT DRACONES

We are with DATA now where we were with GEOGRAPHY *circa* 1510 (and they thought they were pretty good as well!)

Hunt-Lenox Globe

via

The Map Myth of Here be Dragons

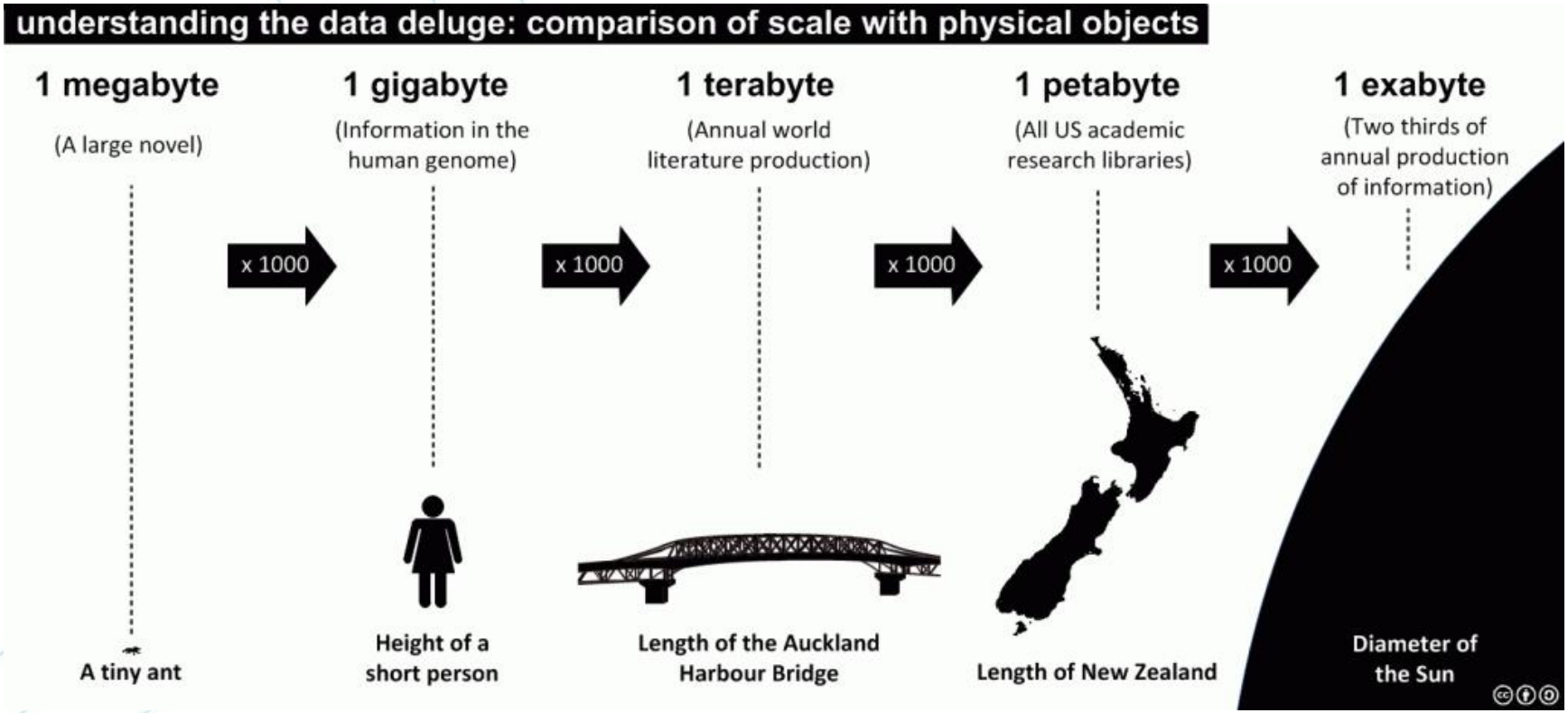
Now for the EVOLUTION
(and the EVIDENCE)...
Knowledge Doubling Curve

Buckminster Fuller created the "Knowledge Doubling Curve"

- He noticed that until 1900 human knowledge doubled approximately ***every century***.
- By the end of World War II knowledge was doubling ***every 25 years***.
- Today things are not as simple as different types of knowledge have different rates of growth. For example, nanotechnology knowledge is doubling every two years and ***clinical knowledge every 18 months***.
- But on average human knowledge is doubling ***every 13 months***.
- IBM states that the build out of the "internet of things" will lead to the doubling of knowledge ***every 12 hours*****!**

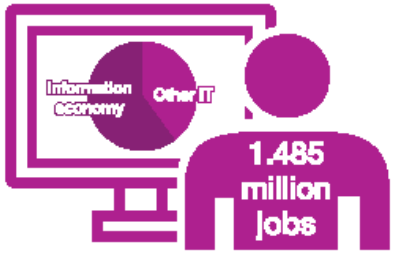
The Data Deluge!

Linear to Exponential Growth of Human Knowledge



HM Government

Information Economy

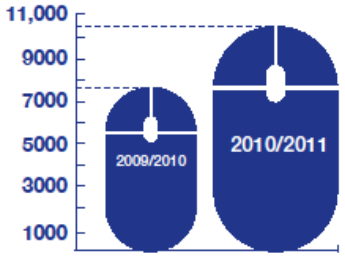


Together there were **1.485 million jobs** in the information economy and IT jobs outside the information economy in 2011

Internet retail volumes grew by more than 6 times between 2003 and 2012, from £4.8bn to £31.1bn

The information economy sector contributed around **£58bn** to Gross Value Added in 2011 (at current prices)

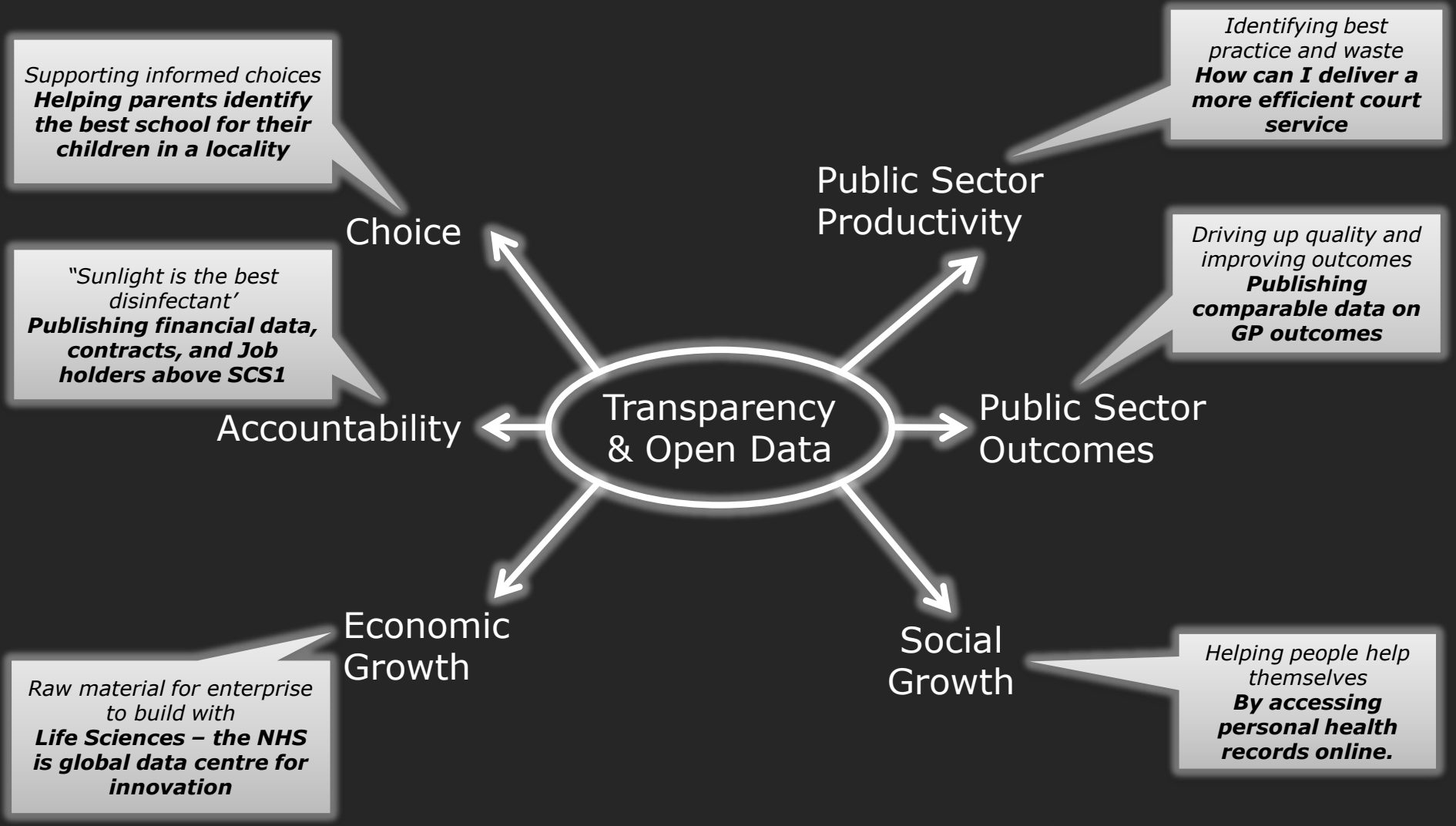
There was a **35% increase** in information and communication technology **apprenticeships achieved** in 2010/11 compared with the year before: 10,510 completed in 2010/11 compared with 7,770 in 2009/10



#infoeconomy #indstrategy

Beyond the data dump ...

Many reasons to open data



Open Data in the UK

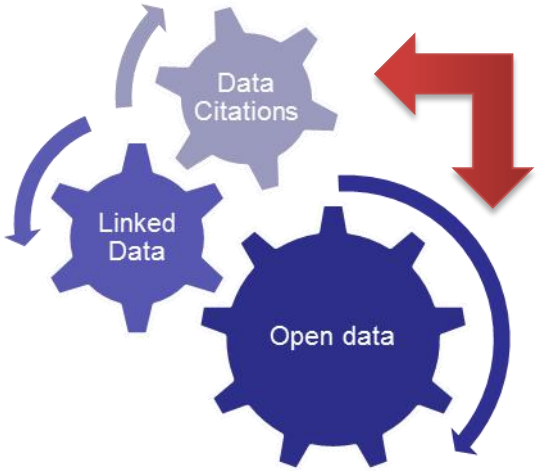
- UK has the most coherent and far reaching Open Data and Transparency strategy anywhere in the World - OFFICIAL!
- UK has some of the world's most comprehensive datasets
- Deloitte analysis as part of the Shakespeare Review* suggested
  - a figure of £1.8bn on the direct economic benefit from use of PSI and
  - a figure of £6.8bn if broader economic and social impacts are taken into account
- Shakespeare states the need for a UK Data Strategy and National Core Reference Data (including marine!)
- Open Data by DEFAULT - Defra's Open Data (Organisation) Maturity Model (about people and systems and ownership and responsibility about data NOT about data *per se*]

* https://www.gov.uk/government/publications/shakespeare-review-of-public-sector-information

And This Means…

- Eventually ("soon") all publicly funded data in the UK will be **Open Data** and **Linkable**
  - This includes marine research and public sector funded marine data... and that's a lot of marine data!
- This makes a bigger base for the DIKW Pyramid
- A lot of rather painful (and costly) encounters with the Dragons of Reality
- NOT ***Caveat emptor*** (although Open Data isn't necessarily free)
- BUT ***imeo Danaos et dona ferentes*** (I fear the Danaans [Greeks], even those bearing gifts)

Linkable Data



Part of Sir Tim Berners-Lee's original vision of the Web was that it should also be used to publish, share and link data. This aspect of Sir Tim's original vision has gained a lot of momentum over the last few years and has seen the emergence of the Linked Data Web.

In realising the vision were moving from Data Islands to the Sea of Linked Data. '*Just as hyperlinks in the classic Web connect documents into a single global information space, Linked Data enables links to be set between items in different data sources and therefore connect these sources into a single global data space. The use of Web standards and a common data model make it possible to implement generic applications that operate over the complete data space. This is the essence of Linked Data.*' Tom Heath and Christian Bizer (2011) Linked Data: Evolving the Web into a Global Data Space (1st edition).

κρατήστε τα άλογά σας
(Hold your Horses)

- Before we go any further, we need to get **Data Management** right and as soon as possible!

WHY?

- Because if you don't know what you've got, what state it is in, how reliable it is, its limitations or even where which version is...
- the potentially bad "bits" get a whole lot worse and quickly ***for data AND data products or services...***
- so we need to develop Data Management Plans

The Role of Data Management Plans

- Government-wide emphasis on community access to data supports a substantive push towards **more open sharing of research data**
- National Science Foundation (NSF) in USA requires **data management plans** as a pre-cursor to it providing Federal funds
- This is consistent with NSF's mission and US policymakers in making sure that any data obtained with federal funds **be made accessible to the general public**
- NSF is subject to the **US Federal Open Government Directive** to make government more transparent and more participatory
- This addresses the trends and needs in the modern era of data-driven science

Biological Database v Biological Journal

…the future dissemination and impact on science?

- If databases and journals remain discrete, our methods of assimilating information will change little in the years to come;
- If databases and journals become more integrated, the way we do science could change significantly;
- The traditional biological journal could become just one part of various biological data resources as the scientific knowledge in published papers is stored and used more like a database and
- conversely the scientific literature could seamlessly provide annotation of records in the biological databases.

Scientific Citations

- Imagine reading a description of an active site of a biological molecule in a paper, being able to access immediately the atomic coordinates specifically for that active site, and then using a tool to explore the intricate set of hydrogen-bonding interactions described in the paper.
- Not only are the data generated by the experiment immediately available within the context of what is being read, but specific tools for interpreting these data are provided by the journal.

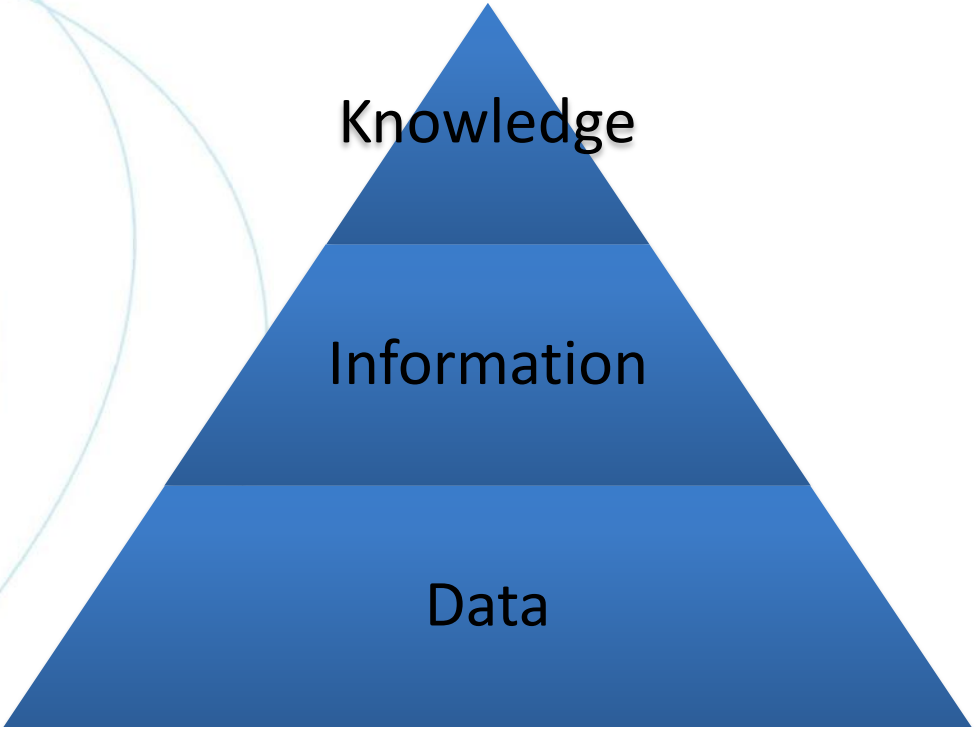
SO CAN WE SUBSTITUTE A MANUAL, BUSINESS OPERATION OR DESKTOP MANAGEMENT DASHBOARD... FOR A JOURNAL?

Yes We Can…

- Not only are the daily operations of databases and journals similar, but the business models also have parallels
- So what is the difference between an entry in a database and an article in a journal?…a mix of perception and content perhaps
- To the user, at least by one measure, the database entry may indeed be more valuable?

http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1193993/

Data - Information - Knowledge - Wisdom

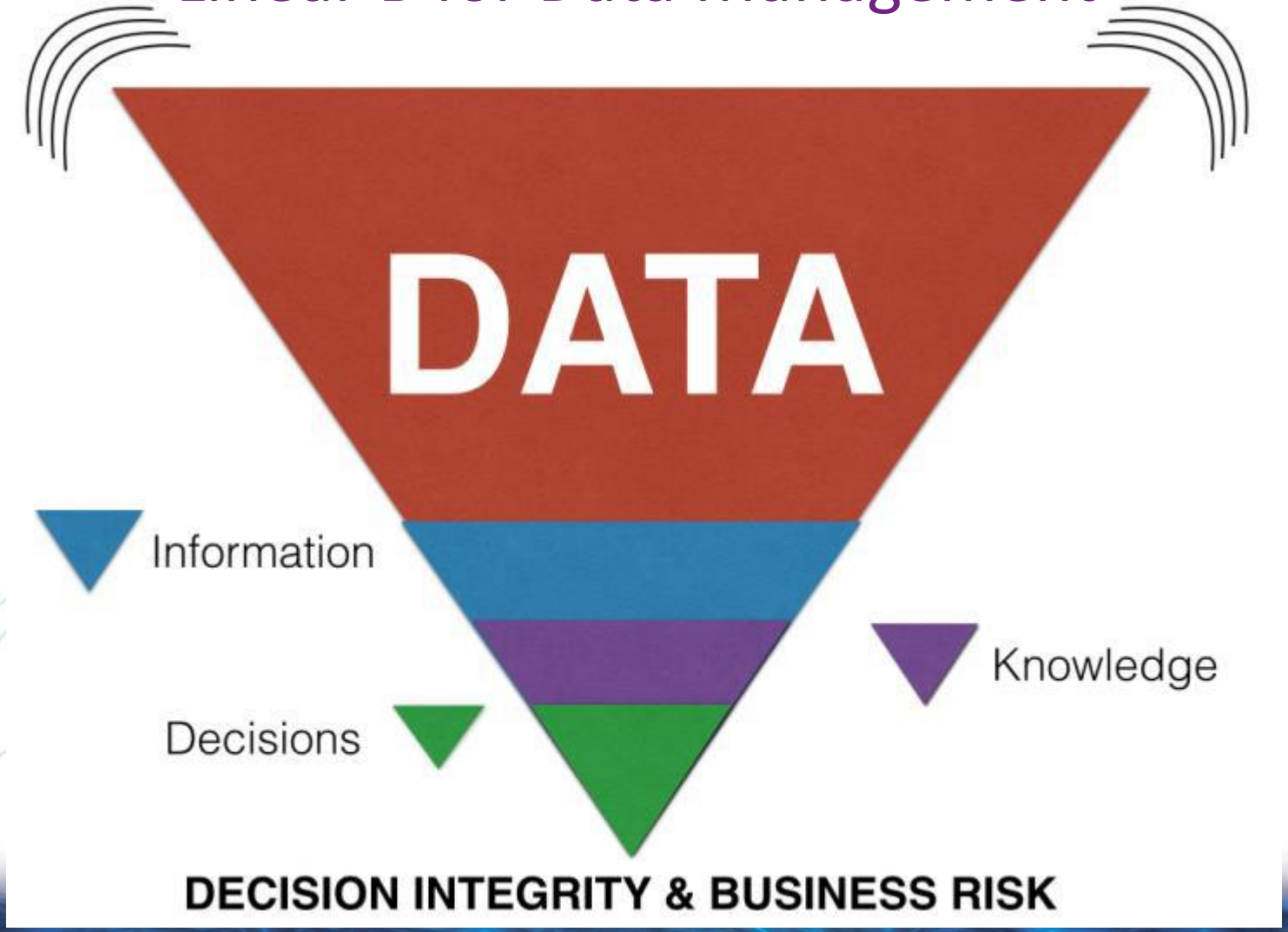


Is this a case of the data at the bottom of the pyramid being more valuable and useable than the knowledge at the top. If so is it time to invert the pyramid?

EVOLUTION says INVERT

REVOLUTION says INVERT

1510 for maps & data
Quill and Double Entry Ledgers for Knowledge
Linear B for Data Management



Many Thanks

John.pepper@oceanwise.eu

David.morris@oceanwise.eu

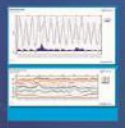
OceanWise

End-to-End Marine and Coastal Data Management and Decision Support

Intelligent Marine and Coastal Mapping Data

Data Policy, Strategy and Management Systems

Enterprise GIS and Productivity Tools



Environmental Data Sharing and Publishing

Capacity Building, Training and Mentoring

www.oceanwise.eu

Independent • Experienced • Innovative • International • Trusted