

Data – A Liability or Asset?

The Role of Data Governance and Integration
in turning your Data from a Liability to an Asset

Dr Mike Osborne
OceanWise Ltd.

BIG Data

Five Star Data

**Internet of Things
(IoT)**

Machine Learning

Linked Data

Semantic Web

**Artificial
Intelligence**

Ontology

BIM

Blockchain





- Drowning in terminology? And data?
- Take back control
- But how?

What We Do

- Marine Data Management and Decision Support:
 - Marine and Coastal Mapping Data
 - Enterprise GIS and Maritime Productivity Tools
 - Training, Mentoring and Capacity Building
 - Data Policy, Strategy and Management Systems
 - Environmental Data Sharing and Publishing
- **Data Integration and Interconnectivity**



Is This You?



Is This You?



Is This You?



Is This You?

- No clearly defined system or formal accountability for the definition, production and use of data
- No one is responsible for overseeing data as a cross-functional business asset
- No data policy, data strategy or processes in place for escalating data issues to senior management
- Lack of effective processes associated with using and maintaining data for decision-making
- Weak rules for controlling sensitive data ref. GDPR

Data 'Entropy'

- More common than you might think ...
- Many organisations are in a state of at least partial 'data anarchy'
- Without policies and procedures i.e. a system of data governance – and constant vigilance – data anarchy prevails



A Common Problem

Data exists in silos
within departments

Or

Embedded in
applications

Resulting in -

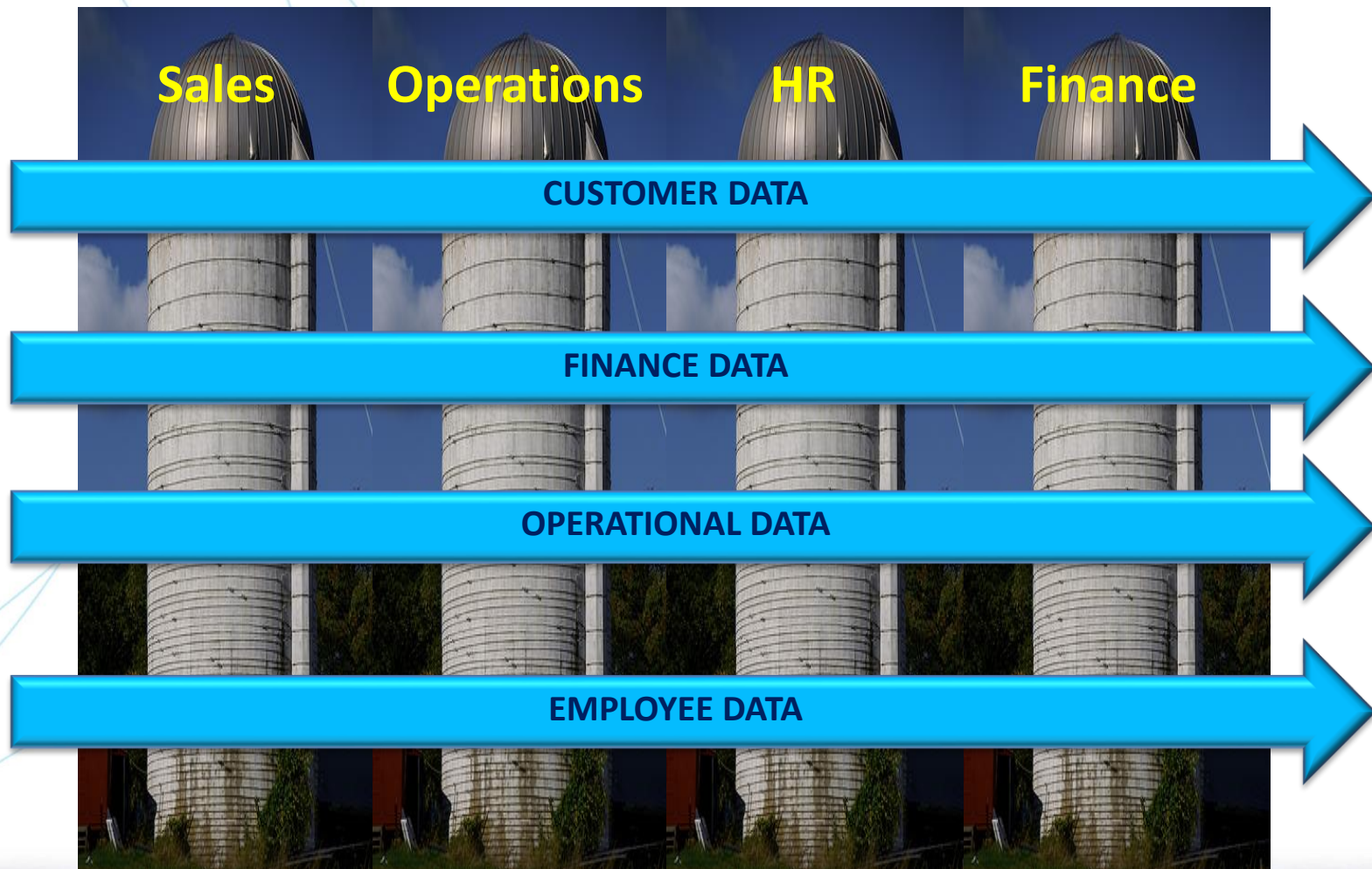
- Inconsistency
- Replication
- Inefficiency
- Confusion

Making data sharing -

- Difficult and
- Time Consuming



Take Back Control Data Centricity - Governance

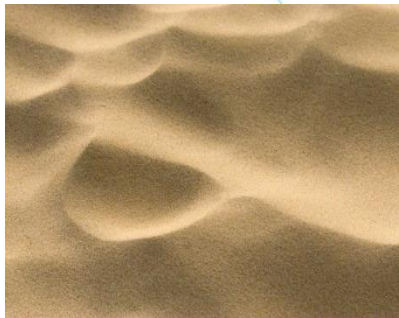


What is Data?

Modern Definition (OED):

- Facts and statistics collected together for reference or analysis
- Things known or assumed as facts, making the basis of reasoning or calculation

Data Volumes



**2.5 QUINTILLION
GRAINS OF SAND
ON EARTH**



**7.5 QUINTILLION
BYTES OF NEW DATA
CREATED EVERY DAY**

90%

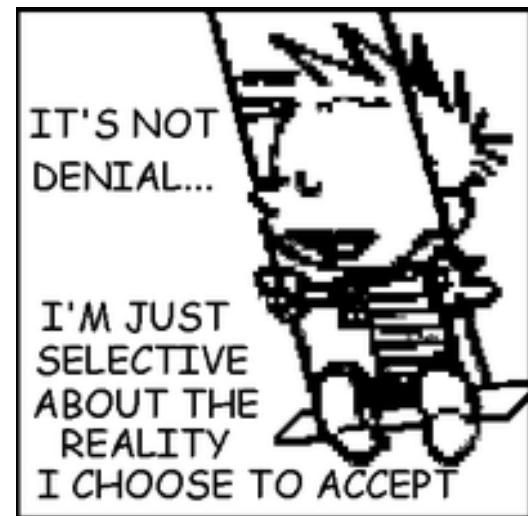
**90% OF ALL DATA HAS BEEN
CREATED IN THE LAST 2 YEARS**

**AVERAGE BUSINESS DATA
VOLUMES DOUBLE EVERY
1.2 YEARS**



Who is Responsible?

- Chief Executive?
- IT/IS Manager?
- Heads of Department?
- Application Provider?
- Nobody?
- Everybody?



What is Data Management?

The development, execution and supervision of plans, policies, programs and practices that control, protect, deliver and enhance the value of data and information assets – in other words:

- What have we got?
- Where do we get it?
- How is it used?
- When is it updated?
- How reliable is it?
- How is it controlled?

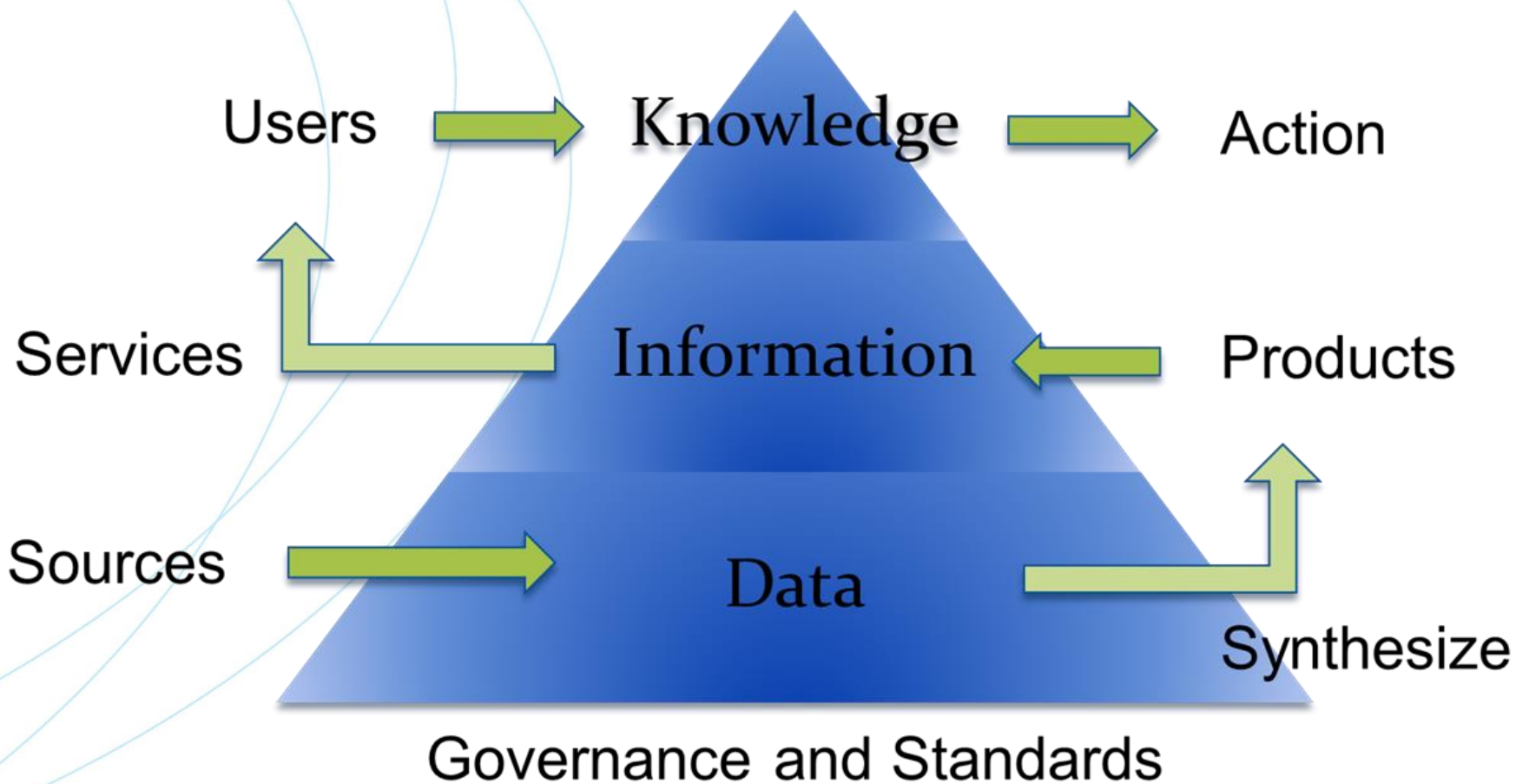


Business Management

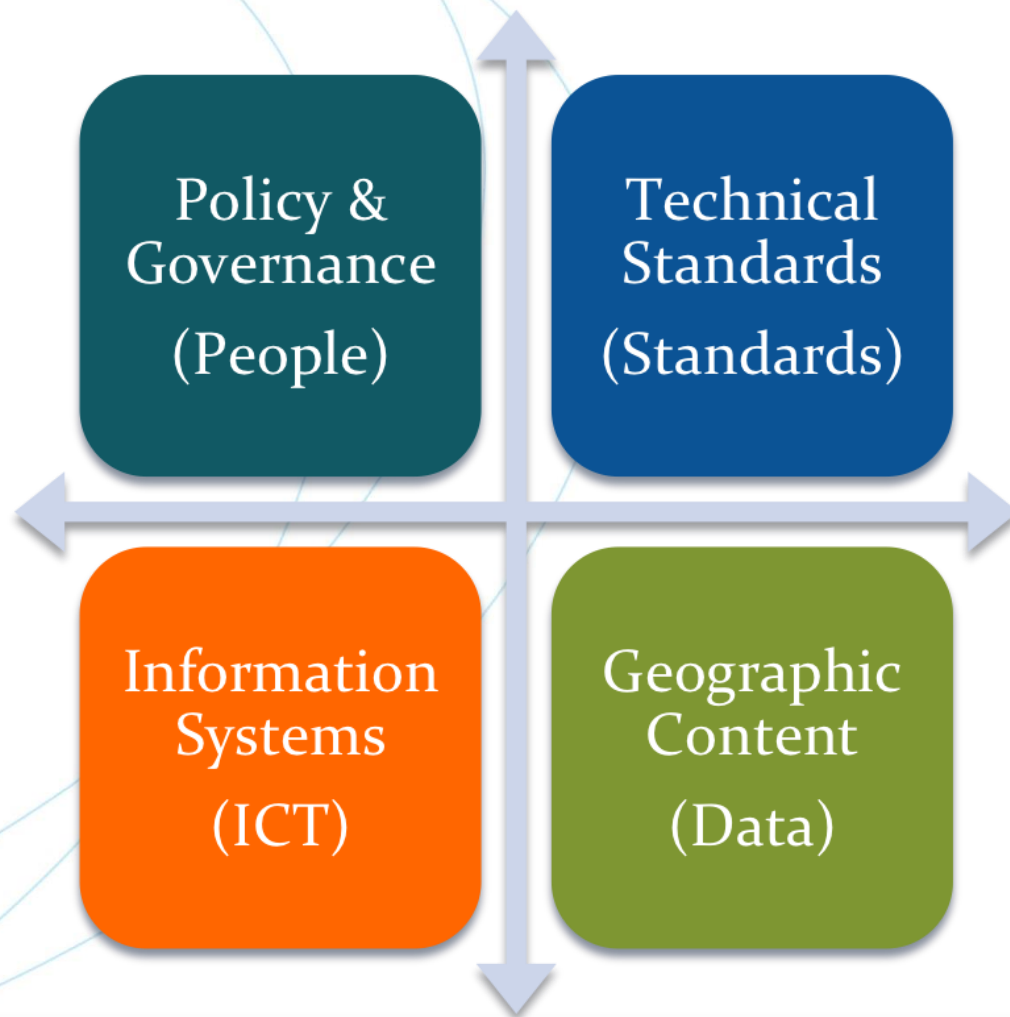
- Policies and Systems for
 - Quality ✓
 - Environment ✓
 - Health & Safety ✓
 - So why not Data?



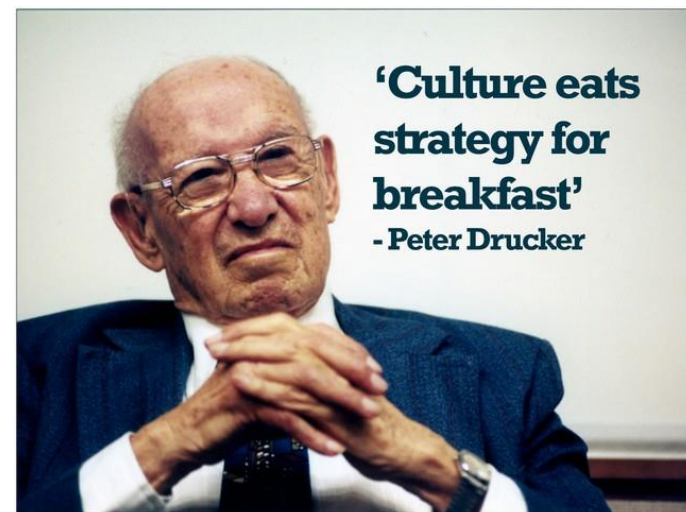
Some Definitions



An Information Infrastructure



Where's
the biggest
challenge?



Some (Data) Principles

1. **Uniquely Identify Data to Facilitate Discovery and Linking**
2. **Create a Master Data Register and Metadata**
3. **Assign Data Stewards**
4. **Acquire Reference Data from Bona Fide Sources**
5. **Implement Consistent Naming Structures for Folders and Files**
5. **Keep Data Accurate and Up-to-date**
6. **Manage Data Close to Source**
7. **Avoid Replication**
8. **Use Standard Reference Frames and a Consistent Means of Transformation**
9. **Use Standard Vocabularies**
10. **Communicate the Above to All Stakeholders**

Data Governance

Definition:

A **process** for **managing** and **improving** data for the **benefit** of all stakeholders



Data Governance

- Data Governance is the execution and enforcement of authority over the management of data and data-related resources
- Data Anarchy is defined as no governance at all
- Organisations that cannot execute and enforce authority over the management of data are most likely in a semi-anarchy state
- Data Governance needs to be communicated and involves internal and external stakeholders

The Role of IT/IS

Is to implement and maintain the technology within an organisation and direct the work of systems and business analysts, developers, support specialists and other computer-related workers

The IT Manager should have an understanding of business and management principles



IT provides the pipework

Data Governance Key Concepts

- Key data items and domains are identified and defined:
 - What are they? (Customer, Supplier, Finance etc.)
 - Where are they are held?
 - Who needs to access them and how?
- Individual business people are made accountable for data within their domain – **Data Stewards**
- All critical data is defined, indexed, measured regularly and reported on by Stewards – **Master Data**
- As problems are identified (reported), initiatives are launched to address them – **Data Improvement**



Where to Start?

- Attend an OceanWise Workshop!
- IMarEST/OceanWise Awareness Course
- And then ?

Do a Data Audit

- What have we got?
- Where is it?
- How is it stored and maintained?
- Who is responsible for it?
- What processes does it support i.e. how is it used?
- Is it replicated and can any of it be centralised?
- Is there scope for streamlining work flows (i.e. to support lean process management)?
- Improvement plan (data strengthening) for legacy data?

Prepare a Data Management Plan

- Implements the Data Policy
- Contains
 - What we will do
 - How we will do it
 - Who is responsible
 - Exchange requirements (what and with whom)
- Covers
 - Acquisition (Data Supply)
 - Management (Storage, Discovery, Access)
 - Publication (Dissemination)

DAMA Framework - Measure Progress

Environmental Factors	+	-	RAG
Vision & Strategy	Strong recognition of the need for DG	No clear alignment between DG and the goals of the organisation	Yellow
Organisation & People	Widespread recognition that ownership of data is required	DG is not seen as business as usual therefore there is a lack of awareness	Yellow
Culture & Communications	Access to shared platforms to help communicate DG messages	No communications plan or ownership of DG communications	Red
Processes & Workflows	Elements of DG methodology in place in parts of the business	No overarching and consistent approach to DG	Yellow
Data Management & Metrics	Some validation of data formats	Insufficient focus on verification of data	Yellow
Tools & Technology	Distributed data sources allow user flexibility and independence	Complex, disjointed and unplanned infrastructure	Red

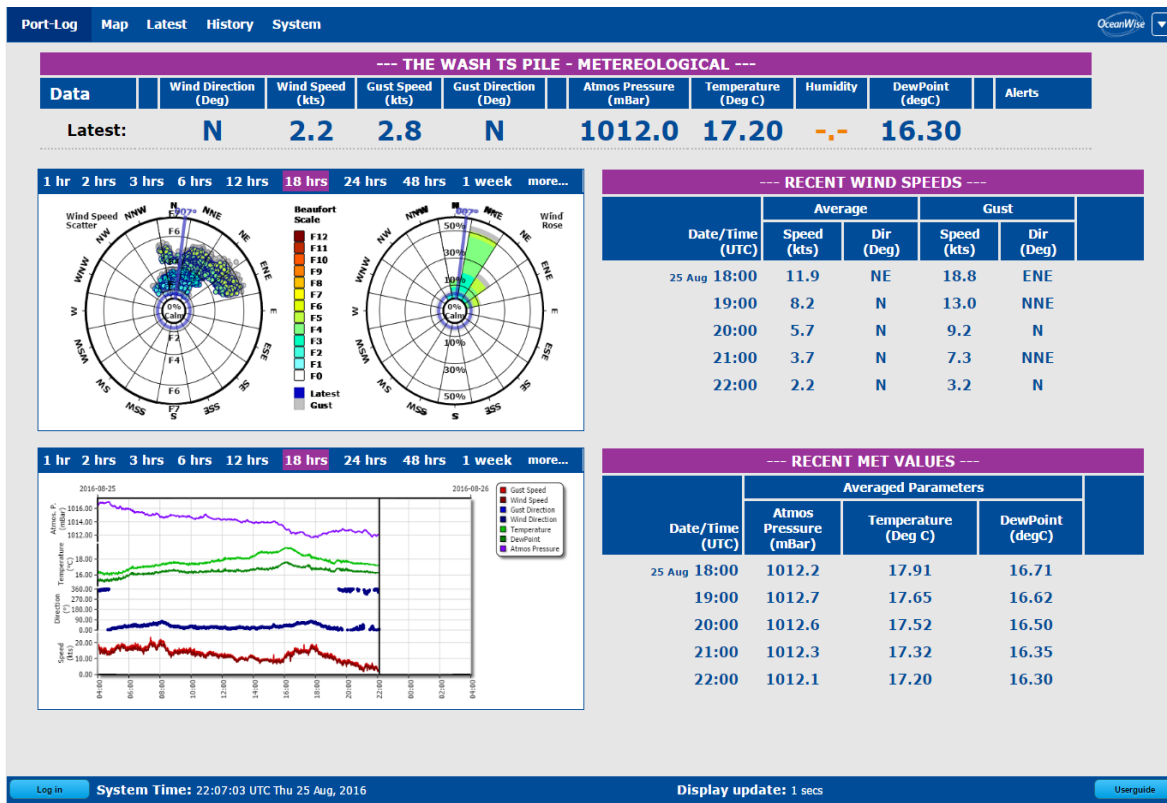
The Result – Common Operating Picture

The data you need easily accessible in one place

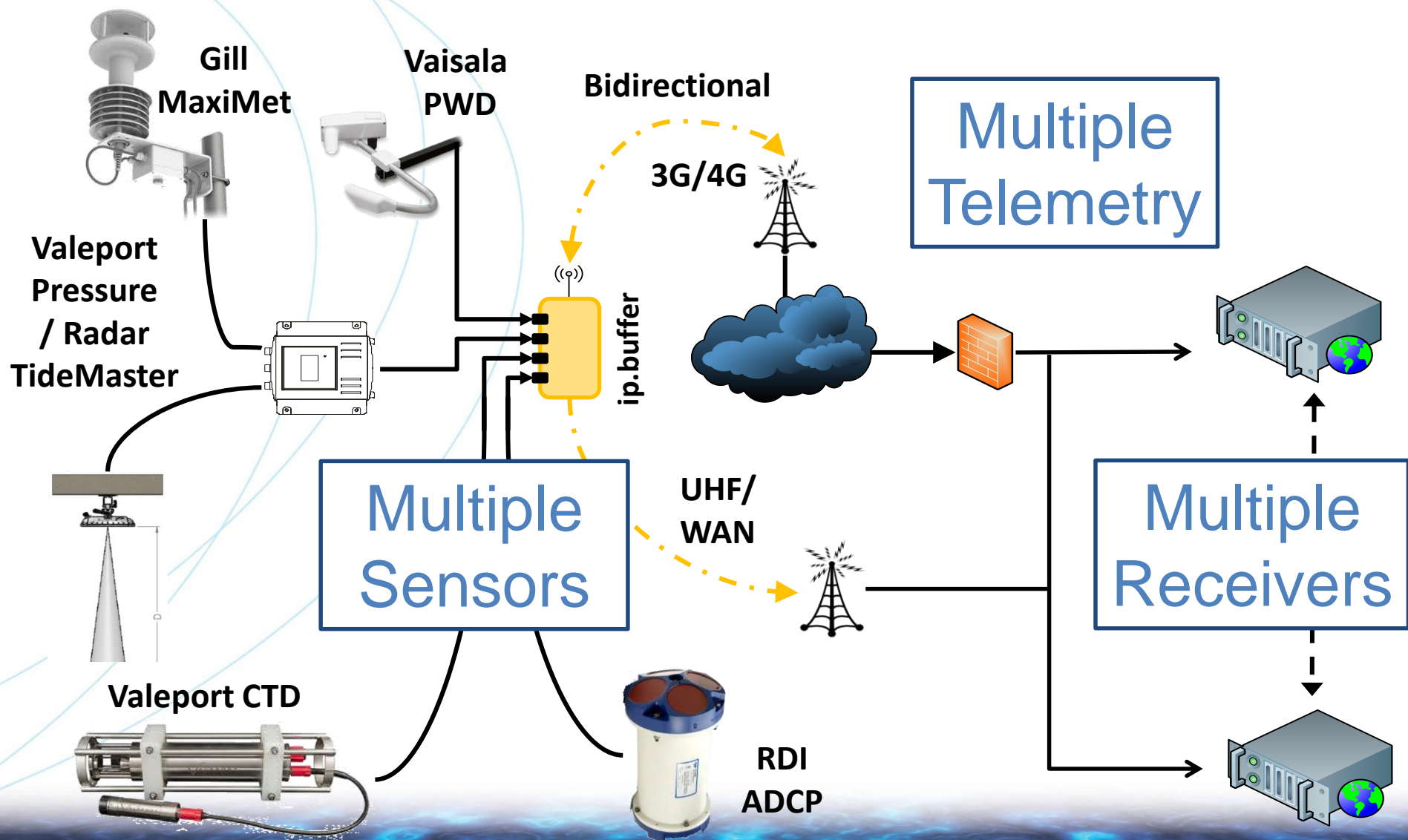


In Space AND Time (An Example)

- Multiple sensors and data types
- Real time and historic datasets
- Data Management and Display System
- Designed and proven in VTS
- Standard and custom web pages

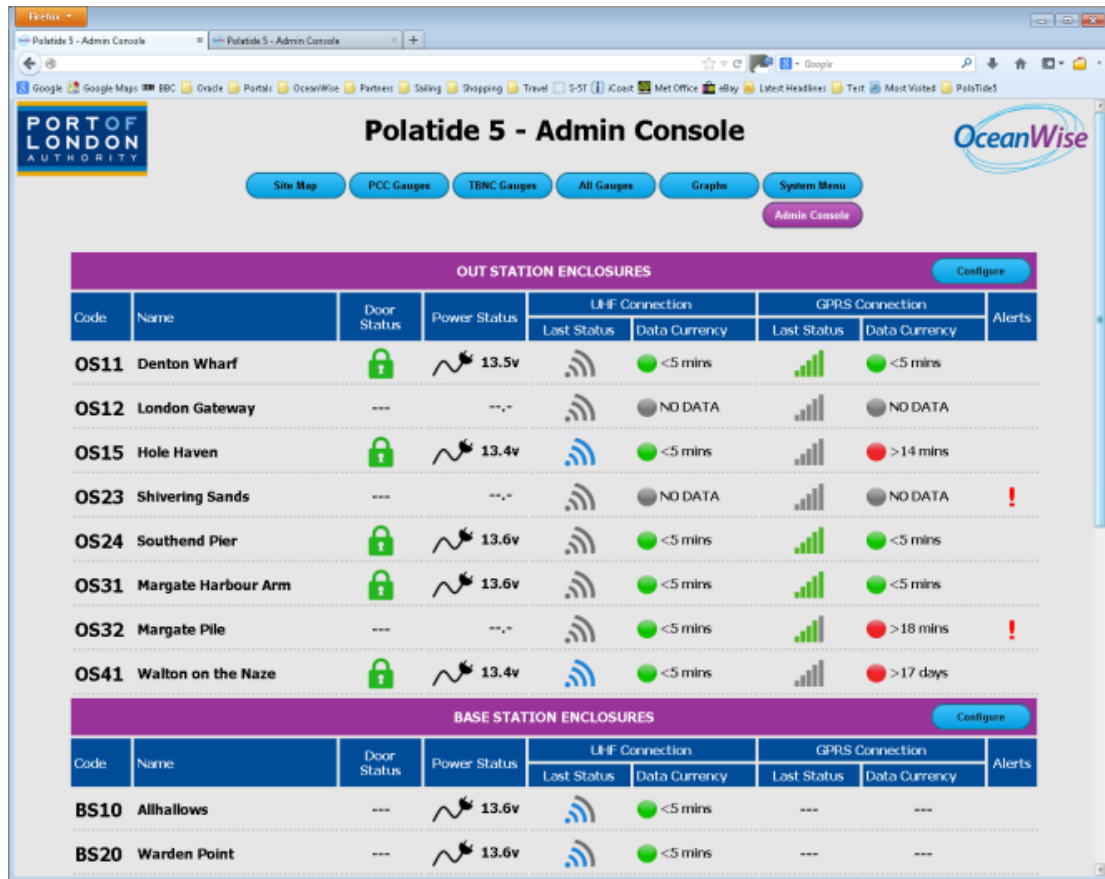


Data Collection and Storage



System Console

- Calibration data and event history
- QC/QA methods and parameters
- System and sensor status
- Alerts and triggers
- User admin.
- Metadata

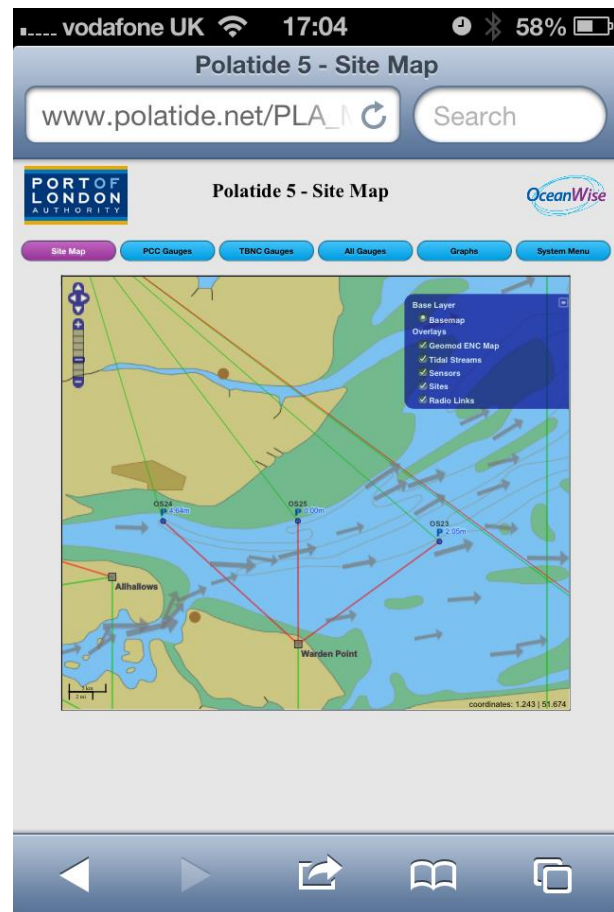


OUT STATION ENCLOSURES								
Code	Name	Door Status	Power Status	UHF Connection		GPRS Connection		Alerts
				Last Status	Data Currency	Last Status	Data Currency	
OS11	Denton Wharf		13.5v		<5 mins		<5 mins	
OS12	London Gateway	---	---		NO DATA		NO DATA	
OS15	Hole Haven		13.4v		<5 mins		>14 mins	
OS23	Shivering Sands	---	---		NO DATA		NO DATA	
OS24	Southend Pier		13.6v		<5 mins		<5 mins	
OS31	Margate Harbour Arm		13.6v		<5 mins		<5 mins	
OS32	Margate Pile	---	---		<5 mins		>18 mins	
OS41	Walton on the Naze		13.4v		<5 mins		>17 days	

BASE STATION ENCLOSURES								
Code	Name	Door Status	Power Status	UHF Connection		GPRS Connection		Alerts
				Last Status	Data Currency	Last Status	Data Currency	
BS10	Allhallows	---	13.6v		<5 mins	---	---	
BS20	Warden Point	---	13.6v		<5 mins	---	---	

Back to Spatial

- User defined base maps
- Application overlays
- Port specific datasets
- Additional functionality
- Scalable screen size
- All good, no Issues
- BUT THEN ...



Please can you interface with

- Company Website
- Public Service Website
- 3rd Party Sensor Network
- AIS Transponder
- AIS Network Controller
- Survey Vessel
- Forecast Provider
- Portable Pilot Units
- Dredgers
- Tunnelling Machine!

Standardisation of course!

- NMEA Sentences
- AIS ASM
- ISO 19100
- S-112
- OGC Sensor Web



**National Marine
Electronics Association**



International Hydrographic Organization
Organisation Hydrographique Internationale



INTERNATIONAL
MARITIME
ORGANIZATION



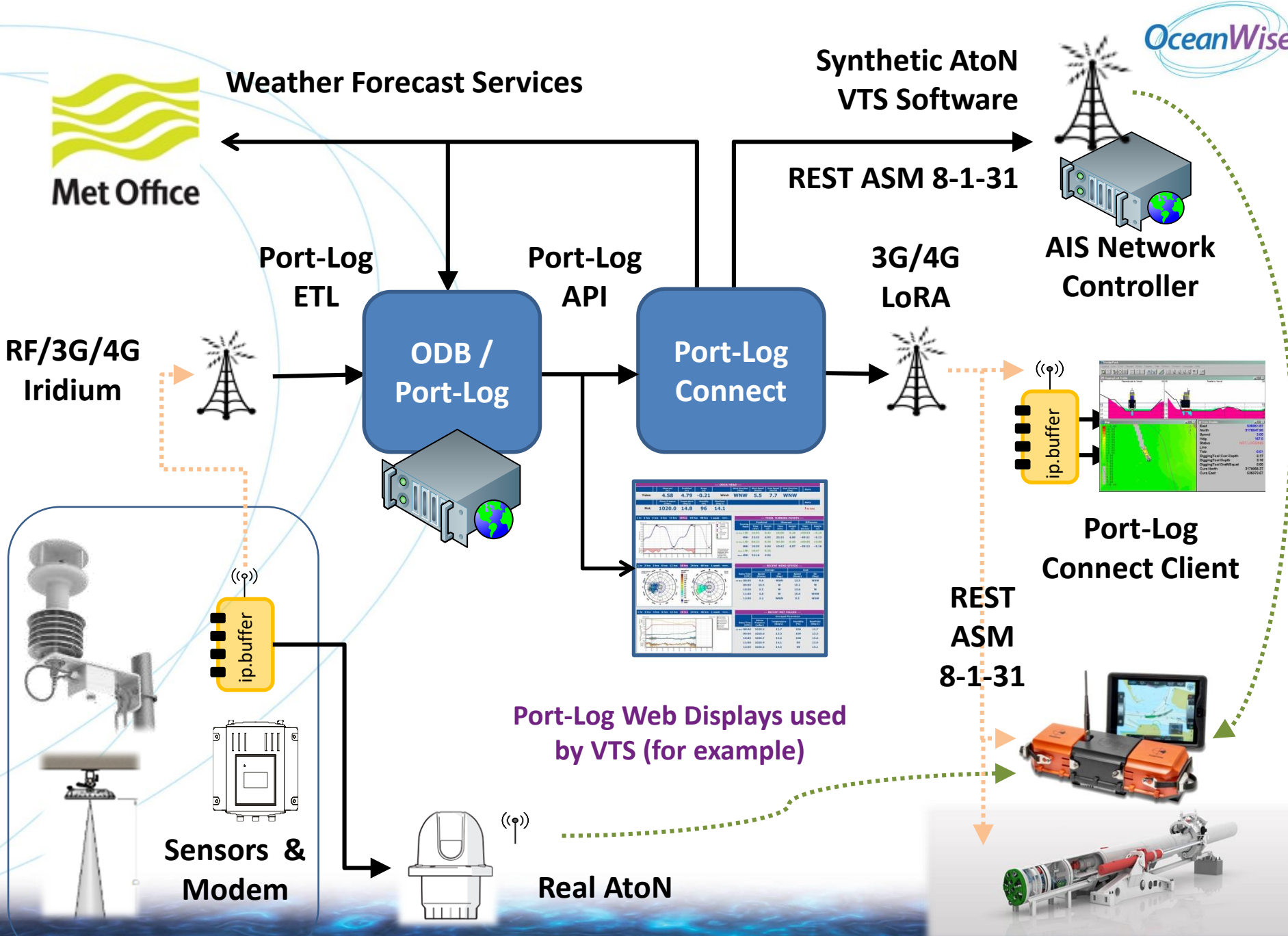
What's available

- NMEA standard / proprietary sentences (Serial)
- AIS Application Specific Messaging (VHF)
- XML (Ethernet etc.)
 - SensorML
 - Other?
- JSON (Ethernet etc.)
 - Sensor Stream Format (SSF)
 - JSON Sensor Format (JSF)
 - Other?

AIS Application Specific Messages (ASM)

Title	Msg	DAC	FI	#slots (max)	State	Registrant
Weather Station Message	6	366	1	5	In force	Saint Lawrence Seaway Development Corporation
Environmental	8	1	26	5	In force	IMO Circ. 289
Meteorological and Hydrographic data	8	1	31	2	In force	IMO Circ. 289
Water levels	8	200	24	1	In force	EU

- A good standard
- Multiple specifications
- Which one to use?
- AIS ASM 8-1-31



Port-Log Real-Time Data

Source	Received	Age	Leader	Data
TW03	16:14:36	6 mins	AIBBM --	IAIBBM,1,1,0,3,8,>jHCv@0Gh0GF2P6uAt3@00HFR06EuOwgl?wnSwe7wvIOwwsAwwnP7mwvh,0*00
TW03	16:14:37	6 mins	AIVDM --	IAIVDM,1,1,,,8>jHCv@0Gh0GF2P6uAt3@00HFR06EuOwgl?wnSwe7wvIOwwsAwwnP7mwvh,0*3D
UDPS 212 227 92 121 6240	16:21:04	1 secs	AIVDM 01	IAIVDM,1,1,,B,13P>ArP05SOrhJLM2TW8?Fb:061D,0*19
UDPS 212 227 92 121 6240	16:21:05	0 secs	AIVDM 03	IAIVDM,1,1,,B,33P8g5@Oh2OrrwpM4?N5SED<00rA,0*62
UDPS 212 227 92 121 6240	16:21:00	5 secs	AIVDM 04	IAIVDM,1,1,,A,402MN7iv6ohDswrrk0M4Eb?020S:,0*2B
UDPS 212 227 92 121 6240	16:21:03	1 secs	AIVDM 05	IAIVDM,2,1,7,B,53P9i0800000hmDd000hmDdU=@E=@0000000000000000Ht0000000000,0*03 IAIVDM,2,2,7,B,000000000000,2*20
UDPS 212 227 92 121 6240	16:19:11	113 secs	AIVDM 08-200-10	IAIVDM,1,1,,A,84hnkj0j2d<<<<<<<0000?'50000,0*57
UDPS 212 227 92 121 6240	16:20:47	17 secs	AIVDM 18	IAIVDM,1,1,,A,B3P;wJ@0?GveugWA
UDPS 212 227 92 121 6240	16:20:24	40 secs	AIVDM 21	IAIVDM,1,1,,B,E>jHC`2W0Q@:7cRa@6400000000OulMR>OhMh:1AACcg@0,4*49
UDPS 212 227 92 121 6240	16:20:49	16 secs	AIVDM 24	IAIVDM,1,1,,B,H3P;wJDU0000000j426p000h5130,0*5C
UDPS 212 227 92 121 6240	16:20:33	31 secs	AIVDM 27	IAIVDM,1,1,,B,K3HhKt1OuKkf;0;@,0*7F
UDPS 77 221 167 70 6908	16:15:53	5 mins	ABVDM --	IABVDM,2,2,1,A,Bh00000000000000,2*04
UDPS 77 221 167 70 6908	16:21:05	0 secs	ABVDM 01	IABVDM,1,1,,B,13cpQg001wOtgRRlMn<9sGv:0H31,0*7B
UDPS 77 221 167 70 6908	16:21:04	1 secs	ABVDM 03	IABVDM,1,1,,B,33P9f0OP0ownrEfM13j=Dgv82>`S,0*53
UDPS 77 221 167 70 6908	16:21:05	--:--	ABVDM 04	IABVDM,1,1,,A,402:oP1v6ohE4woA:HLKNp700p36,0*49
UDPS 77 221 167 70 6908	16:21:04	1 secs	ABVDM 05	IABVDM,2,1,2,B,53P9i0800000hmDd000hmDdU=@E=@0000000000000000Ht0000000,0*0D IABVDM,2,2,2,B,0000000000000000,2*2E
UDPS 77 221 167 70 6908	16:18:27	3 mins	ABVDM 07	IABVDM,1,1,,A,702:oP3dTnnl,0*72
UDPS 77 221 167 70 6908	16:19:11	113 secs	ABVDM 08-200-10	IABVDM,1,1,,A,84hnkj0j2d<<<<<<<0000?'50000,0*5C

Concluding Remarks

- Data volumes are increasing and can't be ignored
- Data exists in silos, is replicated and/or is tied to specific applications
- Port data is unconventional and complex
- Many organisations are in a state of partial anarchy, ill-equipped and struggling to keep up-to-date
- Link data management to business management
- Take back control ...

Thank You!

mike.osborne@oceanwise.eu

