



Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

# Data Quality in the Water Industry

**Presenters:**

**Nicolette Voskuilen, Quentin Griffiths, Greg Preston**

# Why is Data Quality Important?

*Turbidity:*

0.2 vs 0.2



# Metadata – Information About the Data Point

# 0.2

- Units: NTU
- Sample Date: 23/08/2023
- Sample Time: 11:58am NZST
- Temperature: 15.2°C

# 0.2

- Units: NTU
- Sample Date: 23/8/2023
- Sample Time: 7:08
- Temperature: not recorded



# Accuracy or Precision?

# 0.2

- Method: Turbidimeter
- Last calibration date: 21/7/2023
- Meter ID: TURB-1234
- Sample Location: Treated Water Tank Inflow

# 0.20

- Method: Field Sample
- Last calibration date: ?
- Meter ID: ?
- Sample Location: Reservoir



# How do we FAIR?

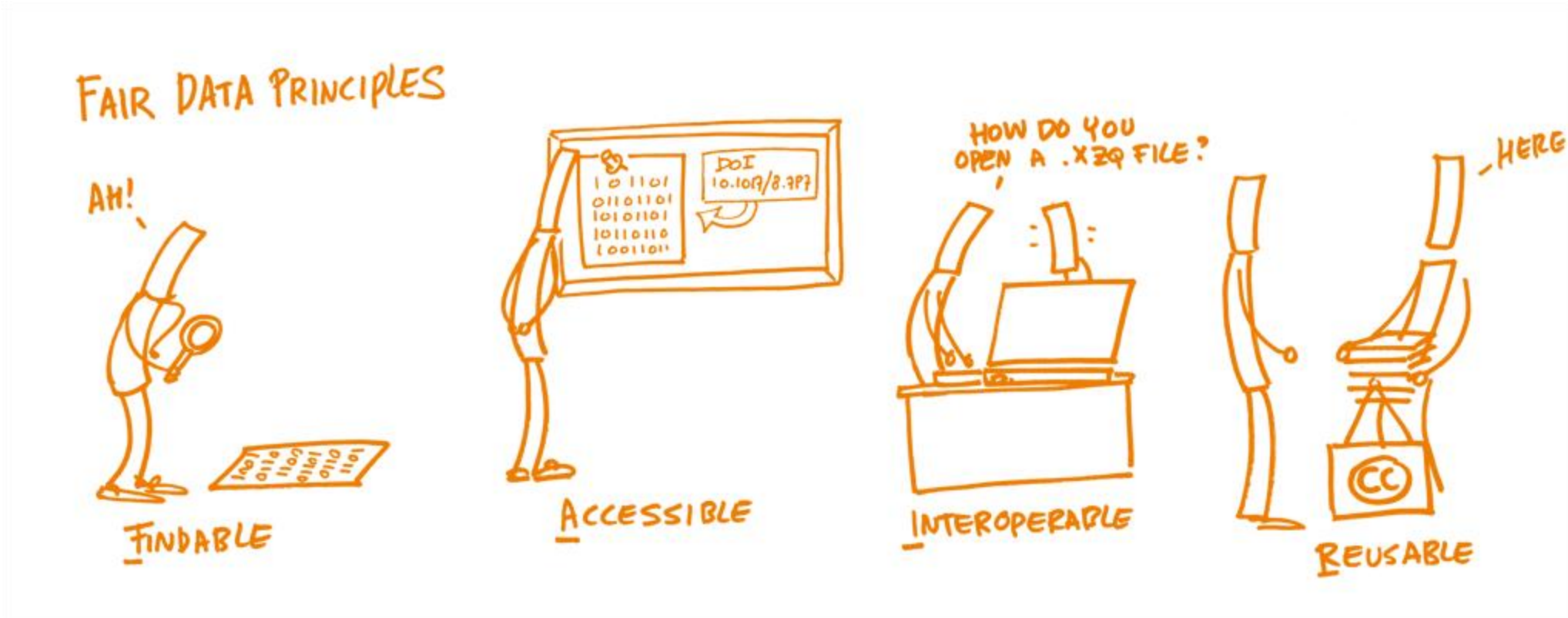


Image Credit: <https://www.fosteropenscience.eu/>



Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

water  
NEW ZEALAND  
The New Zealand Water & Wastes Association Waiora Aotearoa

# How do we FAIR?

## FAIR DATA PRINCIPLES



Image Credit: <https://www.fosteropenscience.eu/>



Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

water  
NEW ZEALAND  
The New Zealand Water & Wastes Association Waiora Aotearoa

# How do we FAIR?

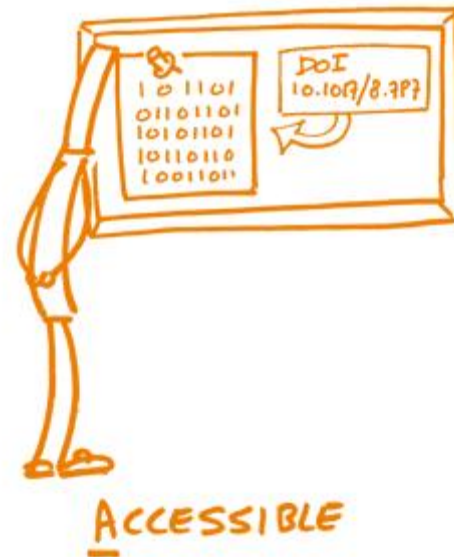


Image Credit: <https://www.fosteropenscience.eu/>



Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

**water**  
NEW ZEALAND  
The New Zealand Water & Wastes Association Waiora Aotearoa

# How do we FAIR?



Image Credit: <https://www.fosteropenscience.eu/>



Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

**water**  
NEW ZEALAND  
The New Zealand Water & Wastes Association Waiora Aotearoa



# How do we FAIR?

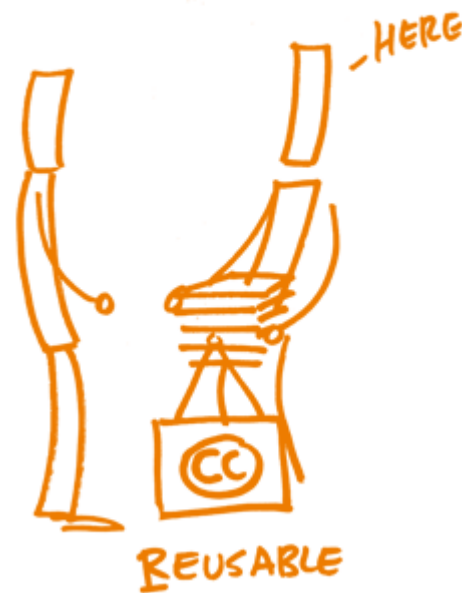


Image Credit: <https://www.fosteropenscience.eu/>



Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

**water**  
NEW ZEALAND  
The New Zealand Water & Wastes Association Waiora Aotearoa

# Data Quality - Underpinning Data Management

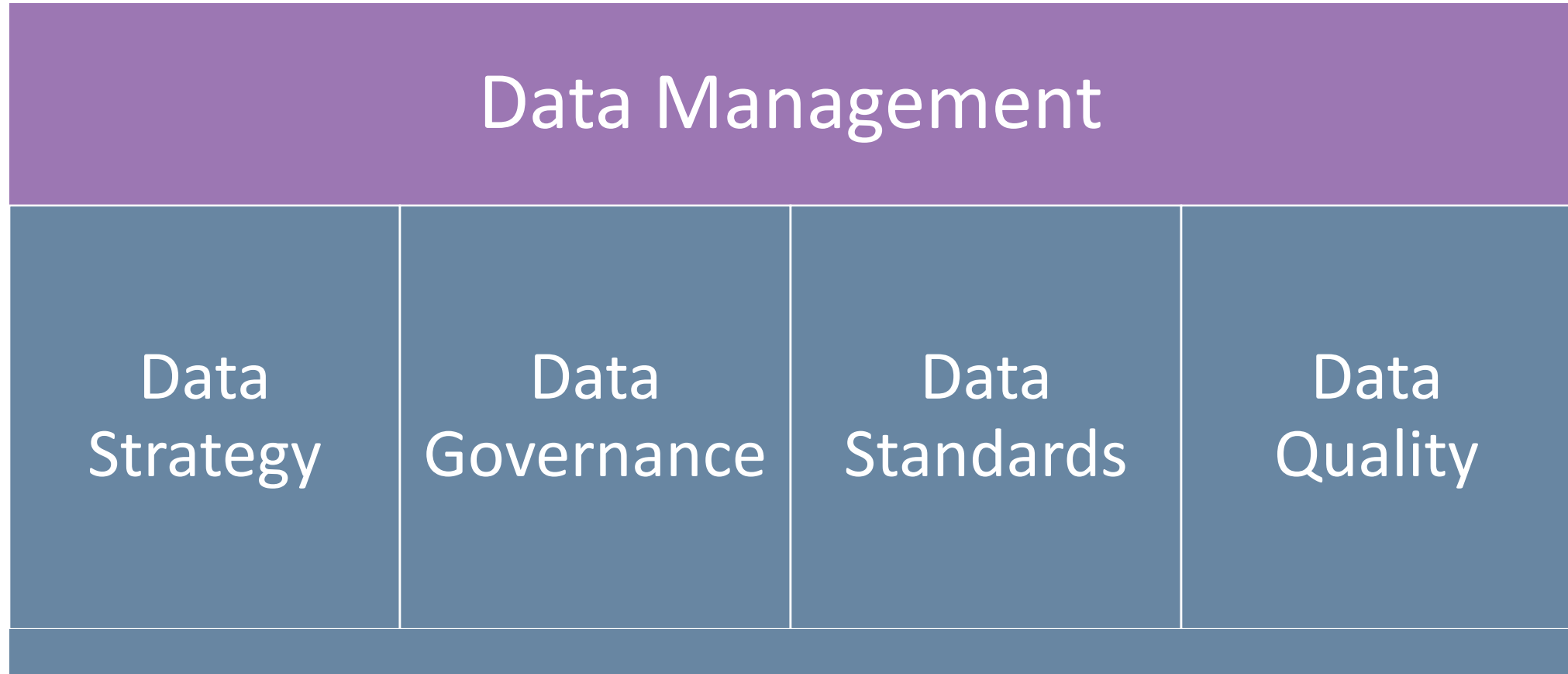


Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

**water**  
NEW ZEALAND  
The New Zealand Water & Wastes Association Waiora Aotearoa

A stylized graphic of three curved lines representing water waves or a mountain range, positioned to the right of the text.

# Four pillars of data management



# Strategy, Governance, Standards

## Data Strategy

- How does the data align with the uses of those data?

## Data Governance

- How do you ensure the processes that enable data, security, integrity, etc?

## Data standards

- What metadata is defined and what format are the data in?



# Data Quality Assessment Framework

Provide a methodology to:

- Assess the quality of data
- Measure and compare the quality of data between different datasets
- Communicate the quality of data in a given dataset



# Data Quality Assessment Framework

Three stages:

1. Data Quality Checks: simple assessments of data with discrete outcomes (e.g. Yes/No)
2. Data Quality Metrics: a quantified aggregation of the results from the Data Quality Checks, designed as interpretable indicators of data quality
3. Data Quality Communication: a set of reporting recommendations to effectively communicate the most important aspects of the quality of data

# 1. Data Quality Checks

Based on Data Standard. Three types:

- **Dataset Level Checks** are carried out on each attribute in a dataset
  - e.g. Does the attribute have data in the same data type as its counterpart in the CoP?
- **Asset Level Checks** are carried out on each attribute, on each asset in a dataset – some are generic, some are attribute-specific
  - e.g. Generic: Is there data recorded in this attribute for this asset?
  - e.g. Specific: For installation dates, is the date a valid date?
- **Geospatial Level Checks** are carried out on a network level for a given pipe asset
  - e.g. Is the pipe in the correct geographic location (within a council's governing region)?



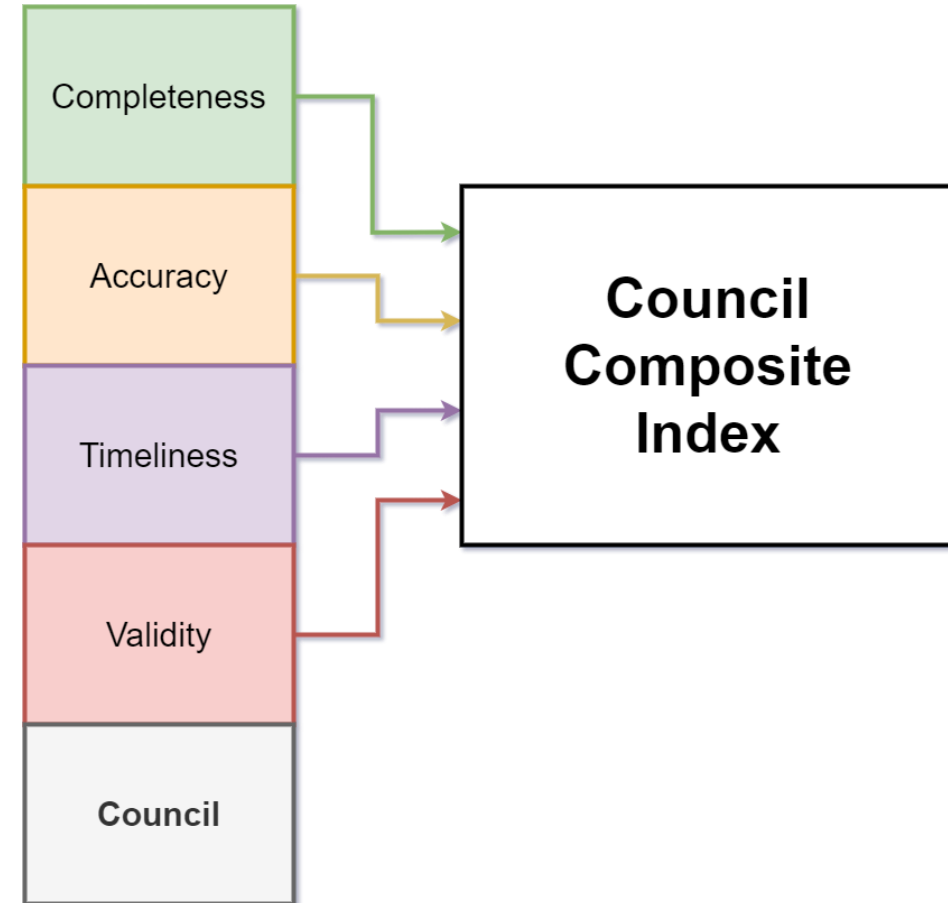
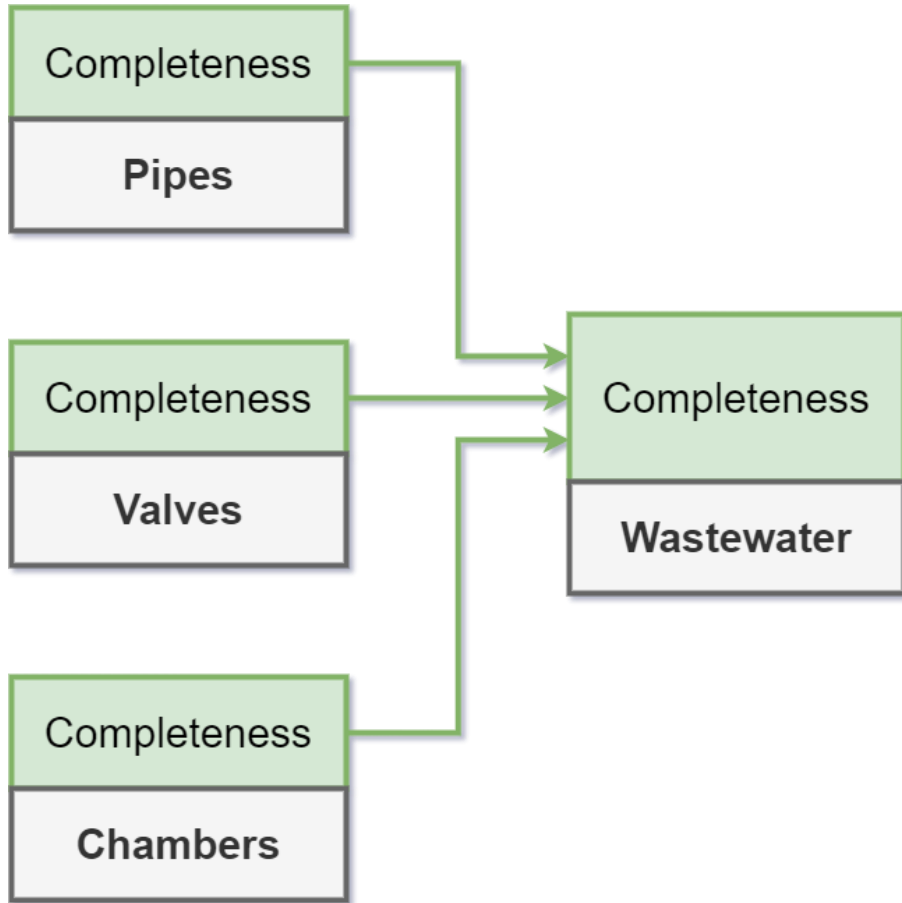
# 2. Data Quality Metrics

Metric	Description
Completeness	The data is comprehensive and does not contain missing values
Accuracy	The data is a true reflection of the real-world values
Validity	The data contains values in a valid type and format
Timeliness	The data is readily available when expected and needed
Data Standard Alignment	The data structure and values are aligned with the appropriate Data Standard





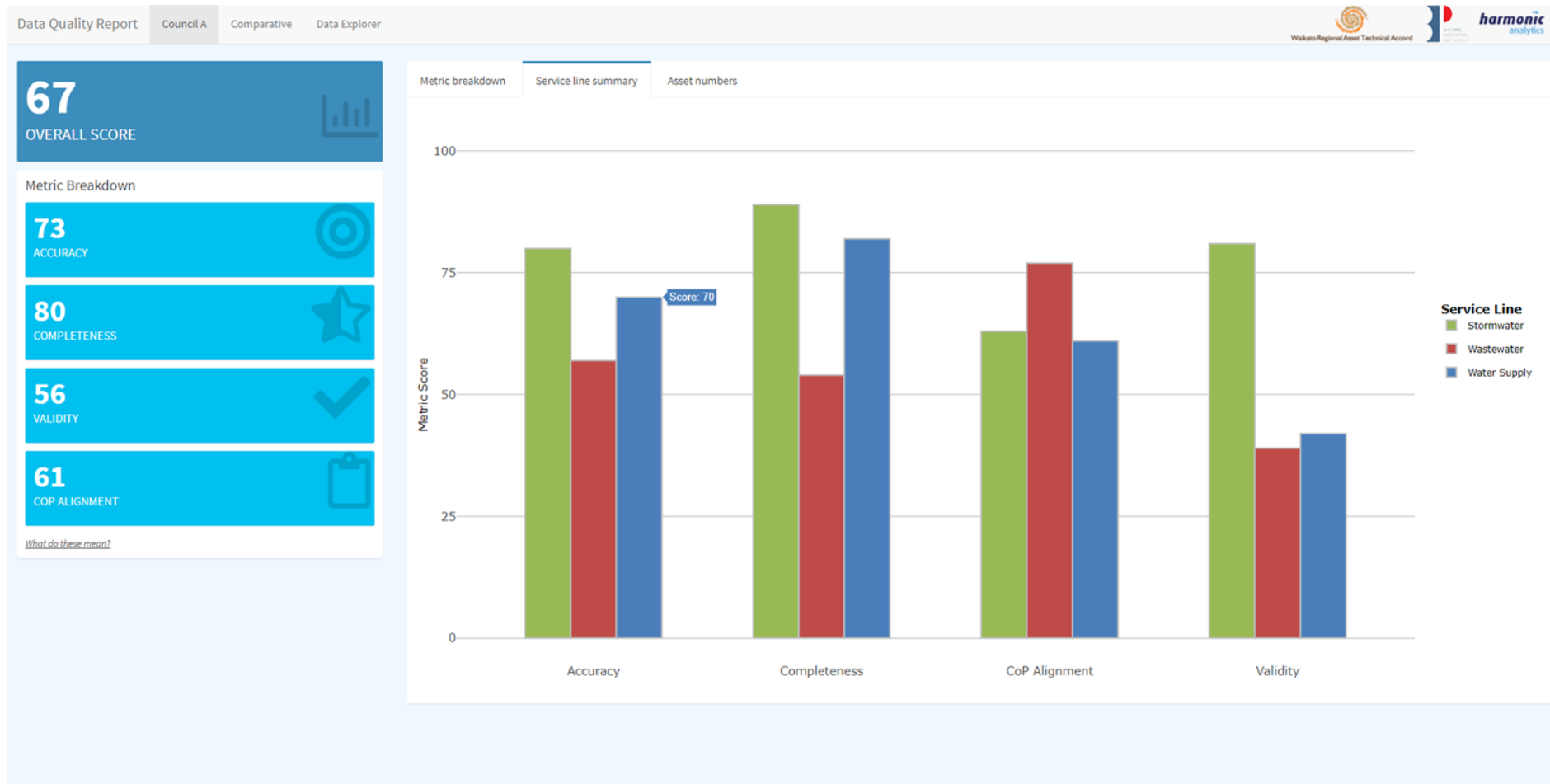
# Data Quality Metrics – Asset Data



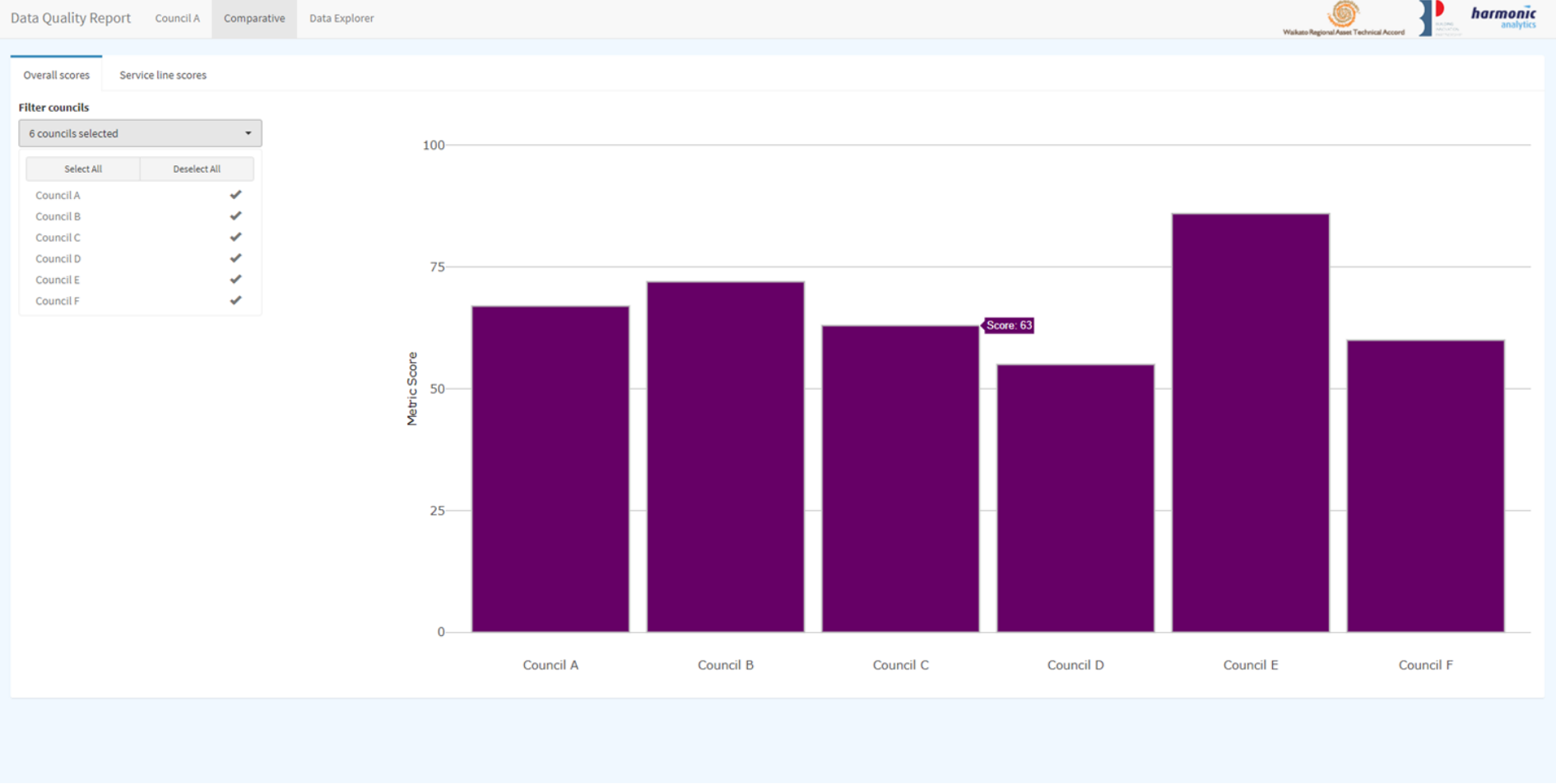
# 3. Communication Dashboard



# Breakdown by service type




# Breakdown by Council



# Breakdown by individual asset

Data Quality Report Council A Comparative Data Explorer

Waikato Regional Asset Technical Accord 

Select dataset to view  
Council - Metric Breakdown  
[Download this data](#)

Show 25 entries

Service Line	Asset Type	Metric Type	Score
Stormwater	Chamber	Accuracy	55
Stormwater	Valve	Accuracy	69
Stormwater	Pipe	Accuracy	32
Water Supply	Chamber	Accuracy	63
Water Supply	Valve	Accuracy	78
Water Supply	Pipe	Accuracy	56
Wastewater	Chamber	Accuracy	39
Wastewater	Valve	Accuracy	82
Wastewater	Pipe	Accuracy	12
Stormwater	Chamber	Completeness	73
Stormwater	Valve	Completeness	23
Stormwater	Pipe	Completeness	65
Water Supply	Chamber	Completeness	50
Water Supply	Valve	Completeness	15
Water Supply	Pipe	Completeness	77
Wastewater	Chamber	Completeness	61
Wastewater	Valve	Completeness	66
Wastewater	Pipe	Completeness	32

Showing 1 to 25 of 36 entries

Previous 1 2 Next



# Conclusion

Useful data is underpinned by:

- A data strategy
- Strong data governance
- Clearly defined data standards
- A robust data quality framework which will:
  - Reference the data standard
  - Set benchmarks for comparison
  - Have a range of aspects including:
    - Completeness, Accuracy, Validity & Timeliness





Smart Water  
Infrastructure Group  
WATER NEW ZEALAND

**Thank you!**  
**Questions? Patai?**