

Regional Council Stocktake

Prepared for Martin Doyle
Tasman District Council

August 2012

Authors/Contributors:

Michael Bargh

For any information regarding this report please contact:

Michael Bargh
Environmental Data Operations Manager
Environmental Information Group
+64-3-343 7806
michael.bargh@niwa.co.nz

National Institute of Water & Atmospheric Research Ltd
10 Kyle Street
Riccarton
Christchurch 8011
PO Box 8602, Riccarton
Christchurch 8440
New Zealand

Phone +64-3-348 8987

Fax +64-3-348 5548

NIWA Client Report No: chc2012-104
Report date: August 2012
NIWA Project: ELF12240

© All rights reserved. This publication may not be reproduced or copied in any form without the permission of the copyright owner(s). Such permission is only to be given in accordance with the terms of the client's contract with NIWA. This copyright extends to all forms of copying and any storage of material in any kind of information retrieval system.

Whilst NIWA has used all reasonable endeavours to ensure that the information contained in this document is accurate, NIWA does not give any express or implied warranty as to the completeness of the information contained herein, or that it will be suitable for any purpose(s) other than those specifically contemplated during the Project or agreed by NIWA and the Client.

Contents

- Executive summary..... 5**
- 1 Introduction 7**
- 2 Methodology..... 8**
- 3 Analysis 10**
 - 3.1 Information managed in which form..... 10
 - 3.2 Metadata history 11
 - 3.3 Is there a metadata standard recognised as the basis of the collected information? 12
 - 3.4 Is a curator role defined for the metadata?..... 13
 - 3.5 How metadata is made available to others. 14
- 4 Summary..... 15**
- 5 Acknowledgements 16**
- 6 References..... 17**

Tables

- Table 1: Council questionnaire responses. 8
- Table 2: Form of information managed. 10
- Table 3: Standards metadata is based on. 13
- Table 4: Summarised standards referenced. 13
- Table 5: How metadata is available. 14

Figures

- Figure 1: How much historic information is maintained?. 12

Reviewed by
Jochen Schmidt

Approved for release by
Jochen Schmidt

J S

J S

Executive summary

It is an intended future that Council environmental information is described and displayed in a consistent manner on a national basis, providing all New Zealanders with the ability to see what data are available; where, why and how it was collected; and what quality it is. Adoption of these practices across New Zealand is promoted by the New Zealand Geospatial Office through the New Zealand Geospatial Strategy.

The purpose of this project was to report how Councils currently manage metadata related to their environmental monitoring installations. In particular we examined what metadata is collected, in what form it is managed, and how it is made available. Information was gathered using a structured 'stocktake' survey that was sent to 15 Councils. 10 Councils responded to the survey with a total of 21 completed questionnaires covering differing environmental monitoring disciplines.

The survey results showed:

- Key deployment meta-information about what is measured (e.g. identifier, parameter, location) is stored in structured digital forms across Councils.
- Environmental meta-information (e.g. site plans and exposure) is often not stored in a structured digital method.
- Other Information such as evidence of data processing, instrumentation maintenance and standard operating procedure are often not stored in easily accessible digital formats.
- Historic information has been collected but it is unclear how much of it is recorded in formats that are not easily accessed through digital interfaces.
- Various standards are used to describe metadata with many being organisationally based. This makes information exchange difficult.
- Curator roles are defined but unclear if there is consistency in tasks.
- Information exchange is done in various, mainly organisationally rather than standards based methods. This is leading to multiple organisations developing similar but different interfaces providing end users with inconsistent methods of access and information.

We recommend:

1. Development of a cross-Council template / standard role description for a curator role as a guideline for Councils. This should include the key tasks.
2. Developing a recommended, system-independent standard for storing deployment metadata using a conceptual framework. Particular attention should be given to those items that are currently weakly managed in digital formats.
3. Development of a standard, system-independent interface for exchange of deployment metadata between systems and to users containing key information required for those exchanges. This could include a list of standard fields (and their

formats) and a set of standard data / service formats (e.g. csv, WFS/XML, shapefile, Excel).

4. Working with relevant system developers in implementing the standards under (3).
5. Promoting the above standards / templates as best-practise through NEMS and other bodies.

These steps will ensure that deployment metadata are managed consistently, independent of agency and/or system and enable easy information sharing across New Zealand.

1 Introduction

It is an intended future that Council environmental information is described and displayed in a consistent manner on a national basis, providing all New Zealanders with the ability to see what data are available; where, why and how it was collected; and what quality it is. Adoption of these practices across New Zealand is promoted by the New Zealand Geospatial Office through the New Zealand Geospatial Strategy.

Currently we know that environmental data held by Councils and other agencies are not universally described, catalogued and 'discoverable' by other organisations. It is often hard for other organisations to know what data are available throughout New Zealand.

The purpose of this project was to report how Councils currently manage metadata related to their environmental monitoring installations. In particular we examined what metadata is collected, in what form it is managed, and how it is made available.

2 Methodology

15 Councils were emailed a 'stocktake' questionnaire in the form of an Excel spreadsheet. They were asked to complete an instance of the questionnaire for each environmental monitoring discipline where they consider metadata is managed in a materially different way. This meant that some Councils returned more than one completed survey.

10 Councils responded to the survey with a total of 21 completed questionnaires. A breakdown of the responses is shown in Table 1

Table 1: Council questionnaire responses.

Organisation	Environmental Monitoring Discipline
Greater Wellington Regional Council	Surface water, Groundwater, Sea level
Greater Wellington Regional Council	Air quality, Meteorological
Greater Wellington Regional Council	Water quality
West Coast Regional Council	Air quality
West Coast Regional Council	Surface water
West Coast Regional Council	Water quality
Horizons Regional Council (Manawatu Whanganui Regional Council)	Surface water, Groundwater, Water quality, Air quality
Taranaki Regional Council	Surface water, Groundwater
Otago Regional Council	Air quality
Otago Regional Council	Surface water, groundwater
Otago Regional Council	Water quality
Environment Southland	Surface water
Environment Southland	Water quality
Northland Regional Council	Surface water, Meteorological, Groundwater
Tasman District Council	Air quality
Tasman District Council	Water quality
Tasman District Council	Surface water, groundwater
Waikato Regional Council	Meteorological, Surface water, Air quality, Sea level, Groundwater, Water quality
Environment Canterbury	Surface water
Environment Canterbury	Air quality
Environment Canterbury	Water quality

The questionnaire was loosely based on the work of Aguilar et al. (2003) and separated metadata into the following categories:

- Station identifiers.
 - Where/how is the source of truth site list maintained? (e.g. site number, site name and aliases, site type, open/close dates, responsible organisation)

- Geographic data.
 - What system is the point of truth for site locations? (including elevation where appropriate)
- Local environment.
 - How is physical site environment information managed? (e.g. site plans, exposure, land use/cover)
- Items/Parameters.
 - What system is used to record the item/parameter list including units?
- Instrumentation and maintenance. How is the asset list managed?
 - Where are calibration records kept?
 - How are fault/maintenance records maintained?
- Data processing
 - What system is used to manage observed data through to its final form? E.g. are there procedures that define permissible changes? Records kept of all data modifications?
- Operational procedures
 - How are operational procedures maintained? E.g. site inspection procedures, instrument calibration procedures, asset maintenance procedures?

For each metadata category the survey respondent was then asked:

- In what form(s) is the information managed?
- Is history of the metadata kept or just current information?
- Is there a metadata standard recognised as the basis of the collected information?
- Is a curator role defined for the metadata?
- How metadata is made available to others.

3 Analysis

Analysis of the survey responses is intended to provide a generalised overview of the current state of metadata collection within Councils. Brief analysis of each metadata question responses is provided in the following sections. The purpose of the analysis is to highlight the general council position related to the metadata question.

3.1 Information managed in which form

Table 2 provides a summary of the form in which information is managed for the metadata categories in the questionnaire. The digitally structured form was separated into two categories; one for where commercial off-the-shelf (COTS) package were referenced, and one for organisational implementations on top of a structured tool, such as Microsoft Excel, Access or other customised database. Digital unstructured forms included electronic documents such as Microsoft Word as well as digital photos.

Questionnaire respondents were able to list multiple answers to the form of management which is why the total responses exceed the total of 21 completed questionnaires.

Table 2: Form of information managed.

	Digital structured - COTS	Digital structured - Organisation	Digital unstructured	Paper
Station identifiers	12	14	7	8
Geographical data	11	12	6	7
Items/parameters	18	5	2	
Data processing	20	6	2	4
Instrumentation and maintenance	8	16	6	13
Local environment	6	9	12	13
Operational procedures			21	

In inspecting this table the general trends that can be observed include:

- Station identifiers, Geographical data, Items/Parameters, and Data Processing are all categories predominantly managed in structured digital forms. Much of the digital unstructured and paper forms are to do with historical information.
- While the Station identifiers and Geographical data categories are predominantly managed using structured digital forms, a bigger percentage of the structured component is made up of custom organisational methods when compared to the Item/Parameters and Data Processing categories. The details in the questionnaires were insufficient to explain this difference but possible reasons include:
 - The COTS solutions were functionally deficient in these areas and organisations pursued custom internal options. There was at least one response indicating a plan to move from a custom implementation to equivalent functionality in their COTS tool now that the functionality existed.

- Organisations are attempting to get reuse in these areas which exceeds the scope of the COTS product.
- The Instrumentation and maintenance category has a higher rate of non structured digital elements than the categories above it. Most of the structured digital elements are custom built, with most COTS support limited to providing instrument comments. The author perceives there is a tension between the need to link instrument information to environmental data and the complexity of coupling this information with enterprise asset management systems. Common practise appears to be to separate the asset management system from the environmental data collection system with manual instruments comment being the coupling method. At least one Council expressed dissatisfaction with this and perhaps it is an area that needs further investigation.
- The Local environment category has a spread with a relatively high rate still not in a digitally structured form.
- Responses in the Operational procedures category were all based on “manuals” oriented at internal operators.

The general assessment of these responses is that the Station identifiers, Geographic Identifiers, Item Parameters and Data Processing metadata categories are generally managed in structured digitally forms. The exception to this is that some historic information which is captured in paper or unstructured digital forms. The predominantly structured digital form of these categories isn't of great surprise as it is this information which is commonly available through web interfaces and supplied as part of any digitally based data exchange.

The remaining metadata categories of Instrumentation and Maintenance, Local environments, and Operational procedures have a higher proportion of the information managed in more unstructured formats. Once again this makes sense as this data has not typically required by external parties. There is an expectation that, in the future, more end users of data are going to want to make their own determination of whether data is suitable for their purpose (including comparison with other data), and therefore are going to want to have access to this sort of information to assist this process. To support this there is a need for further movement of this information into structured digital forms so that it can be made available to others in a consistent and standard way.

3.2 Metadata history

Many environmental data collection sites have already been operating for many years and so a question was added to the survey to see how much historic metadata exists. As Figure 1 shows significant historic metadata has been recorded across all metadata categories.

The survey did not make it possible to determine how much of historic metadata is in paper or unstructured digital forms so it is difficult to assess the size of task that is still faced should this information need to be migrated into a form more suited to making it accessible. It is however good news that the information exists should a business case be developed for its migration to more digitally structured forms.

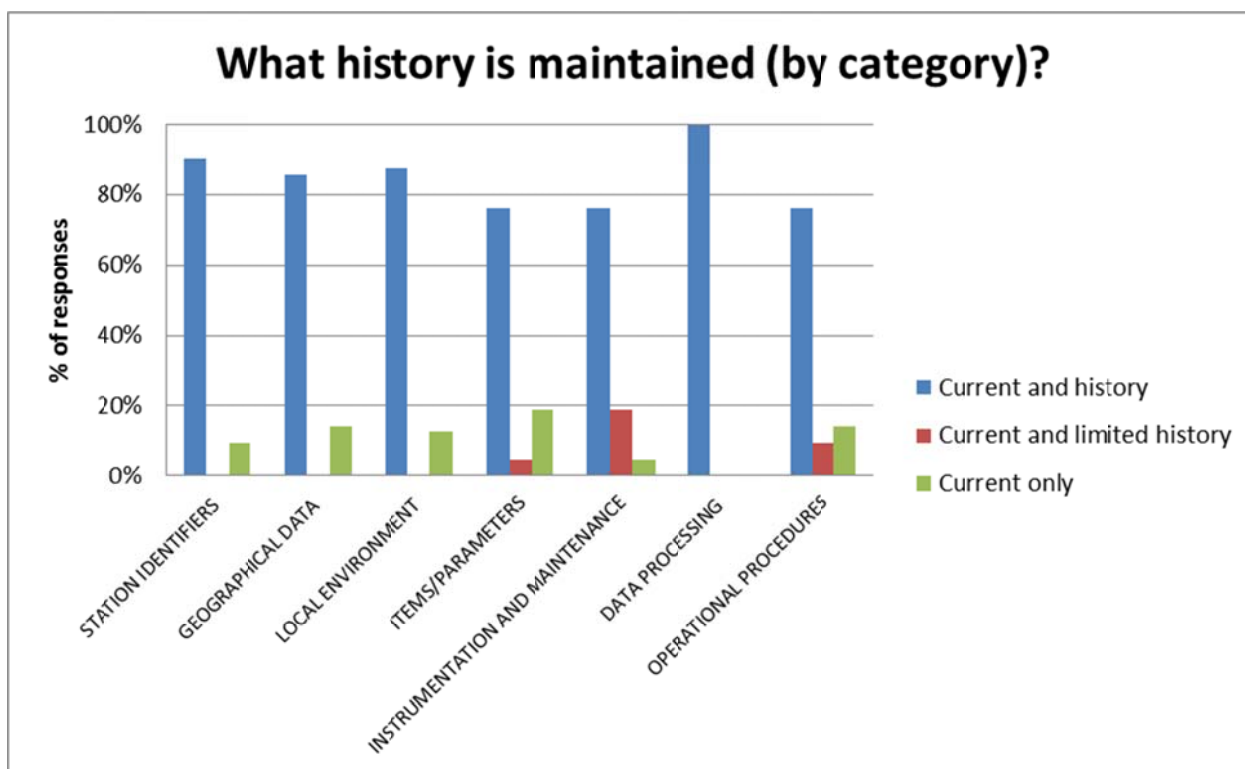


Figure 1: How much historic information is maintained?.

3.3 Is there a metadata standard recognised as the basis of the collected information?

One of the survey questions examined what standards were recognised as the basis for the metadata collected. Questionnaire respondents were asked to describe any standards used against each of the metadata category. Some responses referenced more than one standard and the overall results are summarised results are in Table 3. This table includes the names of referenced standards where they were provided.

No standards referenced in Table 3 are specifically metadata standards. The ISO references are not to the ISO 9115 metadata standard for geographic information but a reference to the ISO9000 based accreditations for quality management systems that several Councils maintain.

The responses shown in Table 3 were further generalised into four categories as shown in Table 4. It shows that 40% responses considered the metadata collected was not based on any standard. A further 27% thought their metadata was based on standards developed by the Ministry of Works and Development, an organisation that disestablished in 1988. 16% of responses based their metadata on organisational standards and this leaves only 18% as based on external standards that may have been developed recently.

The lack of any direct reference to metadata standards in any questionnaire response combined with a low reference rate to any form of recent external national or international standards indicates that there is an opportunity for council wide review of environmental monitoring standards with particular reference to metadata within that. It is noted that there

is a currently active Council led process development environmental standards and it is suggested that it should be checked that metadata standards are linked into this process.

Table 3: Standards metadata is based on.

	Number of references	Percentage of references
No	53	40%
Ministry of Works and Development	27	20%
Organisational	10	7%
Ministry of Works and Development Organisational	8	6%
Organisational, ISO	5	4%
ISO	5	4%
National	4	3%
AS/NZ, US EPA, WMO	4	3%
Manufacturer	4	3%
NZTM	4	3%
NESAQ	3	2%
Organisation, best practise	2	1%
National and organisational	2	1%
No, Ministry of Works and Development	1	1%
Organisational	1	1%
Best practise, organisational	1	1%
Grand total	134	100%

Table 4: Summarised standards referenced.

	Number of references	Percentage of references
No	53	40%
Ministry of Works and Development	36	27
Other external standard	24	18%
Organisational	21	16%
Grand total	134	100%

3.4 Is a curator role defined for the metadata?

Of the 141 category based responses to the question whether a curator role was defined for the metadata there was only 1 response which definitively said there was not. This single response was related to the management of local environment metadata. It is very encouraging to see this recognition that there is a need for organisational roles to ultimately be responsible for the acquisition and care of metadata. A further area for investigation would be to examine what actual activities are current taken by people assigned curator roles.

3.5 How metadata is made available to others.

Responses to the question relating to the availability of metadata to others were varied and it was not always clear (due to a shortcoming of the questionnaire design) what audience had access to the various forms of data indicated. Table 5 is the author's summary of the responses.

Table 5: How metadata is available.

	Internal	External
Station identifiers	Generally available through a COTS or Organisation system	Provided with web views of station data (where available). Limited download capability provided by some councils. More detailed information (and list views) available by request.
Geographical data	Generally available through a COTS or Organisation system	Provided with web views of station data (where available). Limited download capability provided by some councils. More detailed information (and list views) available by request.
Local environment	Generally available through a COTS or Organisation system This includes drive shares and Document Management Systems	One Council indicated a station photo is available via the web but generally this information is only available by request.
Items/parameters	Generally available through a COTS or Organisation system	Provided with web views of station data (where available).). Limited download capability provided by some councils. More detailed information (and list views) available by request.
Instrumentation and maintenance	Generally available through a COTS or Organisation system	No Council indicated this is currently available via the web but not indicated it would be available by request
Data processing	Generally available through a COTS or Organisation system	Survey responses suggests metadata relating to data processing is available by request.
Operational procedures	Generally available through a COTS or Organisation system	No Council indicated this is currently available by the web but not indicated it would be available by request.

Horizons Regional Council's response included a reference to working a Sensor Observation Service (SOS) interface to their system but this was the only reference to a web service based interface that was identified. It was a surprise to the author that this was the only interface mentioned as it was thought there may have been more mention of data transfer interfaces driven by water metering initiatives. It is possible that water metering has led to the development of new transfer interfaces but these were considered to be out of scope of Councils as they do actually undertake the monitoring directly. Whatever the case, the author believes that, across the board, councils will be required to deliver more advanced interfaces to their environmental information and that there is opportunity to overlap this work with interfaces for water metering initiatives.

4 Summary

The survey results show:

- Key deployment meta-information about what is measured where (e.g. identifier, parameter, location) is stored in structured digital forms across Councils.
- Environmental meta-information (e.g. site plans and exposure) is often not stored in a structured digital form.
- Other Information such as evidence of data processing, instrumentation maintenance and standard operating procedure (the 'how' of monitoring) are often not stored in easily accessible digital formats.
- Historic information has been collected but it is unclear how much of it is recorded in formats that are not easily accessed through digital interfaces.
- Various standards are used to describe metadata with many being organisationally based. This makes information exchange difficult.
- Curator roles are defined but unclear if there is consistency in tasks.
- Information exchange is done in various, mainly organisationally rather than standards based methods. This is leading to multiple organisations developing similar but different interfaces providing end users with inconsistent methods of access and information.

We recommend:

1. Development of a cross-Council template / standard role description for a curator role as a guideline for Councils. This should include the key tasks.
2. Developing a recommended, system-independent standard for storing deployment metadata using a conceptual framework. Particular attention should be given to those items that are currently weakly managed in digital formats.
3. Development of a standard, system-independent interface for exchange of deployment metadata between systems and to users containing key information required for those exchanges. This could include a list of standard fields (and their formats) and a set of standard data / service formats (e.g. csv, WFS/XML, shapefile, Excel).
4. Working with relevant Regional Council system developers in implementing the standards under (3).
5. Promoting the above standards / templates as best-practise through NEMS and other bodies.

These steps will ensure that deployment metadata are managed consistently, independent of agency and/or system and enable easy information sharing across New Zealand.

5 Acknowledgements

The author would like to acknowledge the following people and organisation for providing responses to the stocktake questionnaire:

- Andrew Cotter, Taranaki Regional Council
- Brent Watson, Horizons Regional Council
- Chris Jenkins, Environment Southland
- Dale Hansen, Northland Regional Council
- Doug Stewart, Waikato Regional Council
- Evelyn Bruhn, Otago Regional Council
- Jon Marks, Greater Wellington Regional Council
- Martin Doyle, Tasman District Council
- Michele Stevenson, Environment Canterbury
- Pete Stevenson, Otago Regional Council
- Stefan Beaumont, West Coast Regional Council
- Teresa Aberkane, Environment Canterbury
- Tony Gray, Environment Canterbury
- Tracey Frame, Otago Regional Council

Glenn Ellery from Environment Bay of Plenty also provided a survey response but it was received just after the report was written. The response has been added to the project directory so that it is available for any further work on this dataset.

6 References

Aguilar, E.; Auer, I.; Brunet, M.; Peterson, T.C.; Wieringa, J. (2003). Guidelines on climate Metadata and Homogenization. WMO/TD No. 1186. World Meteorological.