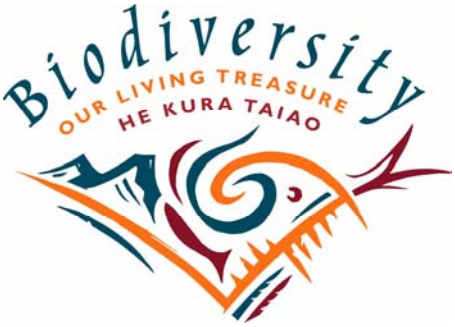

A Survey of Regional Council Biodatabases

Prepared for the Terrestrial and Freshwater Biodiversity Information System Fund by Jerry Cooper, Landcare Research and Julian Carver, Seradigm Limited

14 July 2005

Biodiversity Funds
Serviced by:
Department of Conservation
Head Office
P O Box 10 420
Wellington
0800 86 2020
www.biodiverstiy.govt.nz

Table of Contents

Summary

Project and Client

> A local iwi showed interest in freshwater crayfish ("koura") and asked if anyone knew of the location of koura in the Taranaki region. We were able to perform an immediate query of the freshwater database and provide the exact locations of all sites at which freshwater crayfish had been found.
> Chris Fowles, Taranaki Regional Council.

Regional councils collect many kinds of biodata[1] for biodiversity, pest, and resource management purposes. The above anecdote exemplifies the good data management practices that exist within many regional councils, but the challenge now is to move to a situation where data can be compared and combined at a national level. This is happening well with some types of data, but there is still significant progress to be made. This report documents a survey conducted on behalf of the Terrestrial and Freshwater Biodiversity Information System (TFBIS) fund to determine the status of regional council biodatabases, proposed future developments, and ability to integrate with each other, and with major national biodatabase systems.

Methods

All 12 regional councils and four unitary authorities were contacted. Full responses were gathered from nine regional councils and two unitary authorities and represented a good cross section in terms of organisation size and biodiversity values.

In total 34 datasets were surveyed, from large comprehensive systems to specialised records of particular flora or fauna. Datasets were used primarily for biodiversity management, freshwater quality, and pests. Freshwater Quality was the most widespread and often for smaller councils the only biodata they actively manage in any formal sense.

Results

Data standards, quality, and data management practices

About half the datasets surveyed were using some form of agreed data standards. There was fairly widespread adoption of data collection standards but little of national or international data storage or exchange standards. Councils are collecting data in relatively standardised ways but setting their databases up in different ways, quite independently of each other. This means that data may be difficult to combine across different organisations.

The use of external authoritative lists of species and place names was fairly common in the medium-sized and larger councils but uncommon in the smaller councils. While approaches to metadata management varied considerably one common factor was geographical information systems (GIS), which seem to be emerging as a natural 'attractor' for biodiversity metadata. Data quality varied considerably, being highest for Biodiversity Management and Freshwater Quality datasets. Smaller councils were just as likely as medium and large ones to have high-quality data, but those ranked as high quality by smaller councils tended to be Freshwater Quality datasets. There were no instances where data seemed at risk of being lost and almost all datasets were said to have a reasonable or high level of future certainty.

Future plans

For many datasets there were plans for ongoing development involving expansion of scope, quality and usability improvements, and providing more external accessibility over the Internet. Six organisations mentioned plans to develop new databases. Two general trends observed were some smaller councils planning to expand from water quality datasets into pest management and wetlands

[1] See Glossary in section 5.2 for an explanation of this and other scientific/technical terms used in the document

datasets, and medium-sized and larger councils moving to more comprehensive biodiversity management systems covering a wide range of data.

Integration and interoperability with major biodatabases, and with other regional councils.

Respondents said it would be relatively easy to provide basic minimum data to national datasets for just under one-third of datasets surveyed. Most said they could 'copy data over' but not connect in real time. Perceived constraints to data sharing with major biodatabases included data exchange standards, data quality issues, legal/privacy constraints, technical connectivity constraints, and lack of time and money. There were no reported intellectual property barriers.

Almost all councils perceived real benefits in data sharing with major biodatabases. Benefits included regional comparison of biodiversity and management practices, placing data in a national context, assessment of trends or biodiversity health nationally, and improvements in data quality and standards adoption. Councils with 'comprehensive' or biodiversity monitoring databases mentioned more benefits and tended to talk more about the 'national picture'. Almost all councils surveyed said they stored some data in external systems held by CRIs or government departments. NZFFD/FBIS and NVS were the most common, and DOC datasets were mentioned a few times. Small councils seem to be storing only freshwater data in external systems. Large councils do this, and are also all storing terrestrial biodiversity data in external systems. Only two councils reported storing pest data in external systems.

Most councils saw benefit in sharing data concerned with cross-boundary issues with neighbouring regions, including pest management, biosecurity, biodiversity values, and shared ecological districts. All respondents were highly negative about the prospect of sharing hardware infrastructure with other regional councils. Some supported the notion of centralised systems; some believed this would not cater to their unique regional needs. Some small and medium-sized councils could see the benefit in central systems and/or distributed copies of systems in terms of reduced costs or improved data quality.

Small councils weren't aware of any 'reinvention of the wheel' between regional councils and seemed less likely to know what systems other councils had. Some thought that overlap or duplication happened more between regional councils and other agencies. There seemed to be some duplication of effort between the 'comprehensive' or biodiversity management type databases existing or under development. A number commented specifically on the need for more communication and sharing of knowledge about what others are doing.

Conclusions

Four themes emerge from the analysis: the possible maturation of pest management into a more national focus; the importance of ensuring that the trend towards the development of comprehensive and biodiversity monitoring type datasets does not lead to fragmentation; the fact that a 'one size fits all' approach will not work with biodiversity data management; and the role that TFBIS plays in influencing the 'architecture' of the national, highly distributed 'meta-system' that is slowly but surely emerging.

Recommendations for discussion

- Look for opportunities to use GIS systems to federate metadata nationally
- Explore the concept of federation in pest information systems
- Fund communication around the entire life cycle of developing monitoring protocols through to the development of database systems
- Create a 'Community of Practice' for biodata managers.

1. Introduction

1.1 Context

Purpose of the report

This document reports on the results of a survey conducted by Landcare research in 2005 to determine the status of regional council biodatabases in terms of proposed developments, barriers to progress, and ability to integrate with other major biodatabase systems.

The management committee for the Terrestrial and Freshwater Biodiversity Information System (TFBIS) fund wished to investigate some aspects of the status of biodatabases in New Zealand. The purpose of the investigation was to provide the TFBIS committee with information to support its strategic planning and future funding decisions.

A similar investigation was run at the same time into the readiness of major New Zealand biodatabases to integrate with GBIF (the Global Biodiversity Information Facility). This is documented in a separate report.

About TFBIS

The TFBIS Programme supports the conservation of New Zealand's indigenous biodiversity, by increasing awareness of and access to fundamental data and information about terrestrial and freshwater biota and biodiversity. The Programme is one of a suite of new programmes that reflects Government's commitment to achieving the goals of the New Zealand Biodiversity Strategy. The background to TFBIS is as follows.

In February 2000 the Government adopted the New Zealand Biodiversity Strategy (NZBS) to halt the decline in the variety of naturally occurring plants, animals and ecosystems in New Zealand.

In June 2000, the Government announced a funding package of $187 million over 5 years to achieve the goals of the Biodiversity Strategy. This funding has enabled biodiversity management agencies to increase their 'hands on' work programmes, e.g. to manage more threatened species and a wider range of ecosystems, and to initiate other new work. The TFBIS Programme has been allocated $9.6 million (GST inclusive) over 5 years and $2.714 million annually thereafter.

The Department of Conservation (DOC) administers the TFBIS Programme, but it is the Department's view that the Programme is for the benefit of all agencies and organisations that contribute to the management of New Zealand's indigenous biodiversity. More information on TFBIS can be found at http://www.biodiversity.govt.nz/land/nzbs/tfbis/tfbis/index.html.

1.2 Survey scope

For the purposes of this survey 'biodatabases' were defined as those involving terrestrial and/or freshwater biodata as held by regional councils and unitary authorities.

Specifically out of scope for the report were biodatabases held by city councils, TLAs, and non-unitary-authority district councils.

1.3 Survey process

The survey involved:

- The development of a set of survey criteria designed to ascertain what biodatabases exist, what the proposed development paths are for those (and new) biodatabases, what is the current level of ability to integrate with major biodatabases held by central government and major research institutions, and what any barriers are to that integration.
- An assessment of the reported status of the regional council biodatabases against these criteria.

The assessments were done using a combination of written responses, phone interviews, and face-to-face interviews with one or more representatives from the responsible organisations.

The survey did not involve a formal 'audit' of the databases (i.e. physically checking the databases for achievement against assessment criteria). The assessments were based purely on written or verbal responses from participants.

1.4 Responses

All 12 regional councils and four unitary authorities were contacted and given the survey. Full responses were gathered from:

- Auckland Regional Council
- Environment Bay of Plenty
- Environment Canterbury
- Gisborne District Council
- Greater Wellington
- Horizons Regional Council
- Marlborough District Council
- Northland Regional Council
- Otago Regional Council
- Taranaki Regional Council
- West Coast Regional Council

Summary responses were gathered from Nelson City Council, Tasman District Council and Environment Waikato. Environment Southland and Hawke's Bay Regional Council were not able to respond due to staff time pressures.

For Environment Canterbury a response was gathered from one staff member and only represents one dataset out of several held by that organisation. For the remainder (10 organisations) the responses were thorough, often involving the contact person (or in some cases the surveyor) gathering information from a number of different staff. These responses could be seen as fairly comprehensive and representing the large majority of terrestrial and freshwater biodiversity datasets in that organisation. This includes the responses from Auckland Regional Council, Environment Bay of Plenty, Gisborne District Council, Greater Wellington, Horizons Regional Council, Marlborough District Council, Northland Regional Council, Otago Regional Council, Taranaki Regional Council and West Coast Regional Council.

2. Conclusions

From analysis of survey results a number of conclusions can be drawn. These are followed by some recommendations for consideration by the TFBIS committee. These conclusions depend on the analyses in the summary of findings in section 4 but are included at the beginning of the document for those that may not wish to review the findings in detail.

- Of the councils surveyed only 25% had heard of or were familiar with TFBIS. It seems likely that more communication would be necessary for TFBIS to have an impact on activities within regional councils.
- Some types of biodiversity data are more commonly collected, and better managed than others. Freshwater quality data are the most common. This is borne out by the fact that for smaller, less well-resourced councils it is the only biodata they appear to formally manage and/or store in major biodatabases. Where resource management issues such as this are common and have been around for a long time there are well-defined monitoring protocols, associated (peer-reviewed) standards, well-established regional databases and well-connected national database systems. For 'newer' areas of environmental monitoring, protocols and standards are less well established, databases are less standardised, and there is less integration with major biodatabases.
- It seems possible that pest management may be another area soon to mature into a more national focus. It is a clear example where 'the state of your neighbour' has big implications in terms of internal border monitoring. Feeding data into a national picture may provide early-warning systems and best-practice management.
- There is a trend towards the development of comprehensive, and biodiversity monitoring type, datasets. These datasets subsume or aggregate data on significant natural areas, PNAP surveys, priority areas for pest control, historical biodiversity data for the whole region, and site-specific state of the environment monitoring data. Quite a number of councils have embarked on or are on the brink of developing these systems. Data monitoring standards in this area are much less mature than freshwater quality, and seem likely to be far more complex. It is possible that these databases may be developed in isolation and end up incompatible in a national sense.
- It seems likely that a 'one size fits all' approach will not work with biodiversity data management. There are real tensions between regional variation and national-level monitoring and research interests. These include variation in environments, biodiversity values, individual competence and enthusiasm, resource levels, culture, and council priorities. Some areas of biodiversity have a stronger national 'pull' than others; some are more 'ground up'. Centralised systems are likely to work well in some situations; distributed systems are likely to work better in others. A key focus will be to ensure that where distributed systems are likely to emerge they evolve in a 'federated' sense, that is they are able to interconnect and share data, either with their neighbours, or in a national 'clearing house' sense.
- Councils' awareness of other councils' systems and practices is relatively low. Increasing opportunities for communication about biodiversity data management between regional councils is likely to reduce unnecessary reinvention of the wheel, fragmentation, and lack of interoperability.
- A 'meta-system' for all biodata at a national level is slowly but surely emerging. Horizons and Environment Bay of Plenty's use of Landcare Research's plant names web service is an example of this. This 'meta-system' will not evolve as a large centralised model. It will be distributed and loosely connected. Its components will be individual systems in CRIs, government departments and ministries, local authorities, museums and universities. Decisions TFBIS makes have significant impact on the architecture of this 'meta-system'.

3. Recommendations

Based on these conclusions, and other findings from the analysis, a number of recommendations are made. These are simply ideas for discussion and consideration by the TFBIS committee rather than fully formed proposals.

1) Consider the implications of GIS being a strong 'attractor' for metadata. Look for opportunities to federate GIS metadata nationally. Consider linkages with existing LINZ and e-Government activity. It may be that one or two examplar projects could demonstrate the capability for GIS metadata to increase awareness of regional council datasets.
2) Look at the concept of federation in pest information systems. Consider leadership through MAF Biosecurity and/or the Animal Health Board. Encourage GIS-based metastandards through the involvement of national 'consumers' of shared data. Achieve cross-boundary benefits as a side effect of this 'national pull'.
3) Explore ways to fund communication around the entire life cycle of developing monitoring protocols through to the development of database systems (federated or not). Ensure sound linkages between the development of NHMS, regional council systems, and researchers and informatics specialists in CRIs. Use this process to mitigate against the risk of next-generation comprehensive biodiversity management systems becoming fragmented. Encourage adoption of international data storage and exchange standards to increase potential research benefit.
4) Create a 'Community of Practice' for biodata managers. This should include both face-to-face meetings and virtual collaboration infrastructure (email discussion lists and shared workspaces). Encourage 'show & tell' and develop a culture of 'find out what others are doing first'.

4. Summary of Findings

The survey asked questions about:

- **Metadata** – database title, custodian, abstract, purpose, spatial and temporal extent, and approximate numbers of records/volume of data
- **Data Management and Standards** – approach to metadata management, use of data standards, use of data dictionaries, data management policies and/or procedures
- **Accuracy/Quality** – the quality of the data, measures used to document/ensure quality
- **Future Development** – proposed development path for the databases, to what extent the data are continuously being maintained, risk of loss, new databases planned, hardware and software platforms used
- **Integration and Interoperability with Major Biodatabases** – ability to provide or exchange very basic collection or observation data, other fields that may be of benefit to researchers, potential barriers, value of integration, use of any external systems to store data (e.g. NZFFD[2])
- **Integration and Interoperability with other regional councils** – benefits, value in sharing technology infrastructure, examples of reinvention of the wheel

Findings are summarised in the remainder of this section, and are provided in detail in Section 5.4.

4.1 Range of datasets

Respondents listed a wide range of datasets. While varied in size, scope and purpose when considered as a whole, some patterns appeared to emerge. The datasets seemed to fall naturally into a number of categories or dataset 'types' (Table 1). While somewhat arbitrary, and by no means rigorous definitions, these dataset 'types' seem useful in terms of analysing responses to specific questions in this survey.

Table 1 Datasets held by the councils surveyed

Type of dataset	No. in survey
Comprehensive	4
Biodiversity Management	5
Wetlands	3
Freshwater Quality	5
Pests	11
Fauna/Flora	6

The following is a brief explanation and some examples for each dataset type:

- **Comprehensive** – datasets that covered the whole region and included a wide range of data including significant natural areas, PNAP surveys, priority areas for pest control, historical biodiversity data for the whole region, and site-specific state of the environment monitoring

[2] NZ Freshwater Fish Database – http://www.niwa.co.nz/services/nzffd/

data. Examples include Auckland Regional Council Natural Heritage Database, Greater Wellington Key Native Ecosystems (KNE) Geodatabase, Horizons ecoBase

- **Biodiversity Management** – datasets that focused on the management of biodiversity within significant natural areas. Examples include Environment Bay of Plenty Land Resources database, Gisborne District Council PMA database, Horizons Natural Areas Database, Marlborough District Council SIGAreas
- **Wetlands** – datasets that focused specifically on wetlands. Examples include Environment Bay of Plenty Freshwater Wetlands Database, Greater Wellington Wetlands Database
- **Freshwater Quality** – datasets geared towards monitoring and reporting on freshwater quality. Typically these involved some or all of sampling, collection or observations data on macroinvertebrates, plants, algae and fish, and various environmental factors such as water temperatures, riparian cover (presence absence), river hydrology. Examples include: Greater Wellington Fish and Macro-invertebrates, Marlborough District Council QDAS, Otago Regional Council Biomonitoring Database, Taranaki Regional Council Freshwater Biological Database, West Coast Regional Council Macroinveterbrates Spreadsheet
- **Pests** – datasets for the management of pest plant or animal species. Examples include Environment Canterbury Wilding Conifer Database, Greater Wellington Pest Plant extent in regional parks (Parks & Forests), Greater Wellington Biosecurity rat/rodent data, Northland Regional Council Pest Eradication Contracts
- **Fauna/Flora** – datasets involving observations of specific types of organisms, for a range of purposes. Examples include Greater Wellington Bird Count data, Phenology data, and KNE Invertebrate Monitoring, Marlborough District Council Ecodata

A full list of datasets grouped by type is available in Section 5.4 in the Appendix.

Responses have also been analysed in terms of size of regional council. 'Size' is of course a complicated issue. Relative size differs depending on whether one is referring to population, rating base, or land area. There are skews where the region includes a large city, and whether land region in relation to the city is small. For the purposes of analysing the responses in this survey a set of figures collected by MoRST during recent evaluation of the environmental output class are used. These figures list relative total council expenditure, expenditure on research and investigations, % of R&I contracted out, and total number of scientists employed. This in itself makes interesting reading and is included in the appendix. For this analysis only the total council expenditure figure has been used. Councils have been grouped into small, medium and large (Table 2).

Table 2 Councils grouped according to size.

'Size'	Total expenditure	Includes
Large	$60–175 million	Auckland, Wellington, Christchurch
Medium	$20–60 million	Waikato, Bay of Plenty, Hawke's Bay, Horizons
Small	Under $20 million	Southland, Otago, Northland, Taranaki, West Coast, Gisborne, Tasman, Nelson, Marlborough

It is important to note that for the purposes of this analysis unitary authorities have been labelled as 'small' even though their total expenditures would put them in the 'medium' bracket. This is because their 'regional' function typically represents around 10% of their total expenditure, on average around $5 million per annum.

4.2 Data management and standards

A number of questions were asked in relation to data management and standards. These included the approach the organisation took to metadata management, their use of data standards, their use of data dictionaries, and whether or not they had written data management policies and/or procedures.

Metadata management

For each dataset respondents were asked to provide metadata including database title, custodian, abstract, purpose, spatial and temporal extent, and approximate numbers of records/volume of data. They were then asked this question:

In regard to the kind of metadata you've listed above (the who, what, when, where, why for each database) what approach do you take to the management of this metadata? Do you use an authoritative schema across all your databases? If so, how is this stored?

Responses varied quite considerably. Five organisations said they had no metadata management system, and/or they just kept the information in their heads. For the most part these were organisations that only had a very small number of biodiversity-related database systems. Three organisations referred to specific documents, reports or manuals they had which described their systems in some way. A number referred either in passing or in detail to their geospatial (GIS) platforms as repositories for metadata. These respondents all said that their GIS systems had rigorous approaches to metadata and that in order to have the spatial components of their data registered as GIS layers, they were required to provide metadata.

Recording metadata about datasets or documents simply for the purposes of making them easier to locate by others is not an activity that comes naturally to most people, especially when they are busy. Given that GIS systems require metadata to function they seem to be an increasingly powerful 'attractor' for metadata. The fact that they are often presided over by a few very technically competent (and detail oriented) people who insist on high-quality data entry also adds to their usefulness as metadata repositories. This in turn makes datasets easier to locate and access.

Data standards

Data standards are important when combining data from different sources. For each dataset respondents were asked to rate their use of data standards (Table 3), and to list standards used.

Table 3 Criteria for ranking quality of council databases.

Rank	Criteria
5	This database uses one or more recognised international data management standards
4	One or more recognised national data management standards are used in this database
3	Some basic industry sector standards are being used for this database (e.g. we are using a standard agreed upon with a CRI or another regional council).
2	Our own internally defined standards are adhered to for this database
1	No particular standards are kept to for this database

Where organisations gave a response such as 2–3 the lowest number has been used for the purposes of statistical analysis. Greater Wellington made a distinction between data collection standards and data storage standards and provided both rankings for several of their datasets. For the purposes of analysis across all respondents Greater Wellington's data collection standards rankings were used. In all instances Greater Wellington ranked themselves significantly higher for data collection standards than they did for data storage standards. This was representative of their view that while the data was of high quality, it wouldn't be all that easy for other people to make sense of it without specific assistance. Results are shown in Fig. 1.

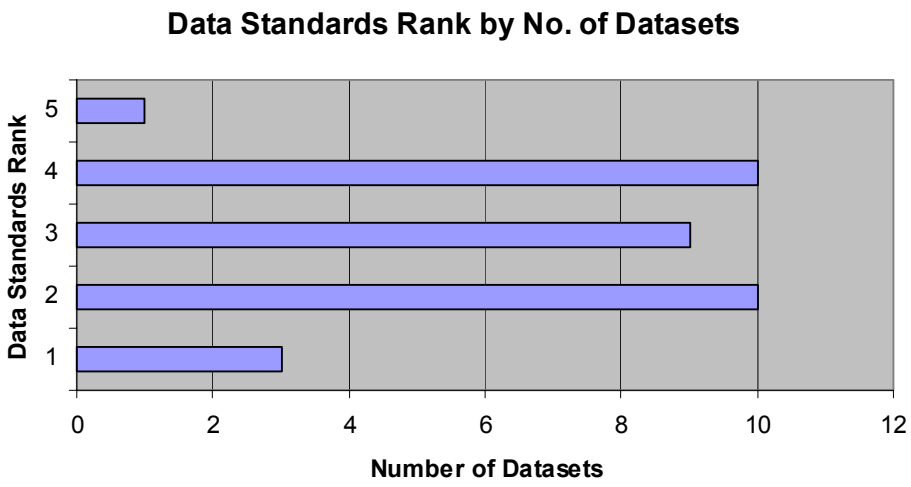

Fig. 1 Datasets ranked by councils according to quality criteria in Table 3.

When measured in terms of total number of datasets, a not unexpected normal curve of distribution emerges. Horizon's ecoBase was the one that ranked a 5, due to their use of the Darwin Core exchange schema standard. From these figures 61% of datasets surveyed scored 3 or above, that is they used some standards external to their organisation. Greater Wellington included a large number of smaller datasets in their response. These ranged in quality from 2 to 4. When these datasets were excluded the total percentage of datasets scoring 3 or above dropped to 47%.

When averaged across the three organisation sizes large councils averaged a 3, medium 2.8 and small 2.7 (Table 4). From these data, as a generalisation it seems that the large and medium size regional councils are more likely to be using national data standards than the smaller ones. Environment Canterbury is an exception to this rule but that seems to have been caused by the inclusion of only one small dataset in their response. Greater Wellington's score is also likely to be affected downwards by their inclusion of a large number of smaller, lower-quality datasets. It should be noted that analysis is fairly subjective due to the variability in the way different people perceive data standards.

Table 4 Average data standards rank by organisation.

Organisation	Total
Auckland Regional Council	4
Environment Bay of Plenty	2
Environment Canterbury	2
Gisborne District Council	4
Greater Wellington	3
Horizons	4
Marlborough District Council	2.6
Northland Regional Council	2
Otago Regional Council	4
Taranaki Regional Council	3
West Coast Regional Council	2

When viewed in terms of dataset type no significant pattern emerged for data standards. Comprehensive and Water Quality datasets rated slightly higher than the others on average.

Where specific industry sector or national standards were listed they included:

- **Biodiveristy Management** – PNAP
- **Comprehensive** – PNAP methodology, SSWI methodology and GRID Wetlands Classification, Darwin Core
- **Fauna/Flora** – DOC protocol, recognised 5-minute and distance count methods
- **Freshwater Quality** – FBIS data standards, MCI (macroinvertebrate community index), national protocol for sampling macroinvertebrates in wadeable streams, NZ Datum, 1947 standard. John Stark biological index for stony streams
- **Pests** – National plant pest recording and reporting guidelines, DOC protocol for collection, entry and storage, NPCA protocol, FORMAK (Peter Handford)
- **Wetlands** – SMF Classification (original version)

When looking at these results and taking into account the conversations had with respondents, some useful conclusions can be drawn. James Lambie's distinction between data collection standards and data storage standards becomes important here. This distinction could be seen as follows:

- **Data collection standards** – recognised methods for gathering, measuring or interpreting data collected in the field. These are typically defined by scientists, emerge through national surveys conducted by organisations like DOC, MfE or AHB, or are influenced by national reporting requirements (such as state of the environment reporting). Examples, as above include MCI, PNAP methodology.
- **Data storage standards** – agreed methods for structuring and naming data in a database. This may include national or international standards for metadata, taxonomic names, geospatial data, or standard schemas for data exchange. Examples include NZGLS, NZ Geospatial Metadata standard, ABCD, and Darwin Core[3].

In regard to data collection standards the majority of organisations appear to be using these where those standards exist. This was especially true where data came from or overlapped with DOC surveys. The standards used seem almost exclusively to be national rather than international. There were, as could be expected, a number of cited instances where organisations had adapted, modified or refined national standards for their own purposes. These were often due to regional variation in the types of environments monitored. Despite this unsurprising regional adaptation and importance of 'local knowledge', staff from regional councils appear to be increasingly collecting data in standardised ways.

When it comes to data storage standards, however, this does not seem to be the case. This has important implications when considering the combination of data from different regional councils. When building their databases to store data collected using national standards, they are structuring these databases in quite different ways. This means that while the data are talking about essentially the same things, it would be reasonably difficult to combine these data across different regional councils.

There are two possible exceptions to the lack of data storage standards adoption. Geospatial data storage standards seem to be increasingly being adopted. It is likely that this is due to standardised tools (e.g. ESRI), the e-government unit and LINZ defining a national standard, and the fact that geospatial data allow the combination of different datasets into 'layers' useful for comparison purposes.

[3] For links to more information about all these standards see the references in section 5.1.

The other exception is Horizon's ecoBase system. When Horizons were developing ecoBase they thought carefully about data storage standards. Says Helmut Jenssen, 'We met several times with DOC's database developers to align ecoBase standards with those developed for DOC's Bioweb. However Bioweb development seemed too complex to accommodate developments outside of DOC, so we continued develop ecoBase on our own. While attending a TFBIS workshop on database integration we had our eyes opened to international data exchange standards. Because of the way we designed ecoBase we found it conformed well with international data exchange standards. This allows us to link ecoBase data with Landcare Research's NVS and Plant Names databases, and ultimately to the Global Biodiversity Information Facility (GBIF).'

Data dictionaries

Data dictionaries are often used to enforce data storage standards in a database. They are sometimes referred to as 'controlled vocabularies' and are often represented as 'pick lists', or 'drop down boxes' in user interfaces. Data dictionaries can be internal to a dataset, or derived from an external source. The following question was asked of respondents:

For the databases you have described to what extent do you use Data Dictionaries as data content standards? Do you use for example:

- *Landcare Research Names Database [4] (for taxonomic plant names)*
- *NZ Geographic Place Names Database[5] (for place names)*
- *Others*

Of the 34 datasets surveyed councils said for 17 they didn't use any data dictionaries, for 8 they used internal data dictionaries, typically of species names or locations. For another 5 they specifically mentioned using lists of taxonomic names from the Landcare Research Plant Names database. Examples of data dictionaries used include:

- **Biodiveristy Management** – Stream names (New Zealand Standard), Plant pest names from National Plant Pest Accord, Landcare Research Plant Names Database, Valuation New Zealand, What's in normal topo sheets
- **Comprehensive** – Landcare Research Plant Names Database, LENZ, GRID wetlands methodology, Link internally to NVS datapoints but not data dictionary style.
- **Freshwater Quality** – Index given to us from Cawthron Institute, Documents on fisheries, changing classification on our native fish in Otago, DOC native recovery programme (non migratory species), Topographical maps for naming river systems
- **Pests** – Tree species of half a dozen tree species (e.g. Douglas-fir, *Pinus*), Surrounding land use (intensive grazing, low-intensity grazing, no grazing), Spread risks on plantation (point scale). Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature
- **Wetlands** – Landcare Research Plant Names Database

Overall Pests and Flora/Fauna type datasets were less likely to use data dictionaries, perhaps because individually they were dealing with a relatively small number of species. With the exception of Gisborne District Council no 'small' councils mentioned external data dictionaries.

During phone interviews, when respondents mentioned internal data dictionaries, they were asked how these were kept up to date. A number replied that they had started with a list from Cawthron, a CRI, or published literature, and then when a new species 'popped up' they added it in manually.

[4] http://nzflora.landcareresearch.co.nz/
[5] http://www.linz.govt.nz/rcs/linz/pub/web/root/core/Placenames/SearchPlaceNames/downloaddataset/index.jsp?

By 'popped up' they referred to themselves discovering new ones, periodically checking against NZFFD, or periodically reviewing the literature. This was true for Marlborough District Council, Taranaki Regional Council, and Otago Regional Council. These responses seemed to be mostly in regard to datasets related to Water Quality. Another pattern that emerged was that the smaller councils, especially those in the west and top of the South Island, tended to outsource macroinvertebrate identification to Cawthron Institute, effectively using them as an ongoing 'data dictionary' or quality-control mechanism.

Another interesting phenomenon was the recent piloting of a plant names web-service provided by Landcare Research to Environment Bay of Plenty and Horizons Regional Council. This allows these organisations to integrate up-to-date taxonomic names directly into their internal systems in real time. This means these organisations will not have to bear the cost, complexity, and risk of errors inherent in maintaining their own lists of taxonomic names. It is an early example of a 'meta-system' emerging out of the integration between many separate systems across different organisations.

Data management policies

Respondents were asked: *Do you have documented data management policies and/or procedures? To what extent are these utilised or complied with?*

For nine out of the 34 datasets surveyed respondents said that they had no documented policies/procedures. For four datasets no response was provided to this question. Comprehensive, Water Quality, and Wetlands datasets had a much higher incidence of documented data-management policies/procedures than the other types of datasets. The larger councils were more likely to have documented policies/procedures than the smaller ones, although often these only applied to one or two major datasets rather than all in that organisation.

Examples of responses in the positive to this question include:

- **Auckland Regional Council** – 'ARC has protocols over management of GIS datalayers and access to information. ARC Information Management is currently being reviewed.'
- **Environment Bay of Plenty** – 'The new Biodiversity databases will use a data depreciation process to manage competing data and metadata will be filled in for all data in the databases. Database administrator will ensure that all data are collected in accordance with standard methodologies.'
- **Greater Wellington, Possum Residual Trap-catch data** – 'ISO 9001:2000 approved Quality Manual, which includes data management policies. Fully complied with (audited by external agent).'
- **Marlborough District Council, QDAS** – 'Documented manuals for data entry processes.'
- **Otago Regional Council, Biomonitoring database** – 'Have a full in-house data management manual. Have a dedicated internal data manager.'

Examples of responses in the negative include

- **Greater Wellington, Fish and Macro-invertebrates** – 'Not known by me, but assumed to have some policies. This Dept explored the idea of an ISO-accredited system but has yet to implement it.'
- **Horizons, ecoBase** – 'Looked into it. ecoBase is a stand-alone Access database, project lined up sometime this year to put it into SQL Server. This will bring it in line with the council's data management processes.'
- **Marlborough District Council, SIGAreas and QDAS** – 'Data entry done only by one or two people so no documented procedures'

- **Gisborne District Council, PMA database** – 'Wouldn't often get changed. Where errors are identified these get changed along with DOC. No formal policies.'

4.3 Accuracy/Quality of data

Having defined-levels of certainty about the quality of biodata is important when those data are used for purposes outside of their original intent. Respondents were given the following definition of data quality:

The OECD Data Access Operating Principle for Quality defines this as:
Quality refers to the proper description of uncertainties surrounding the production of the data (e.g. the techniques employed in their collection and archiving, and the measuring instruments and their calibration), the ability to ensure that the cited source and value are ***authentic***, that the data retain ***integrity*** (complete and absent from introduced errors), and that they are ***secure*** against loss, destruction, modification, and unauthorised access.

Respondents were then asked to rate the kinds of data quality measures taken according to certain criteria (Table 5), and the datasets showed the overall distribution in Fig. 2.

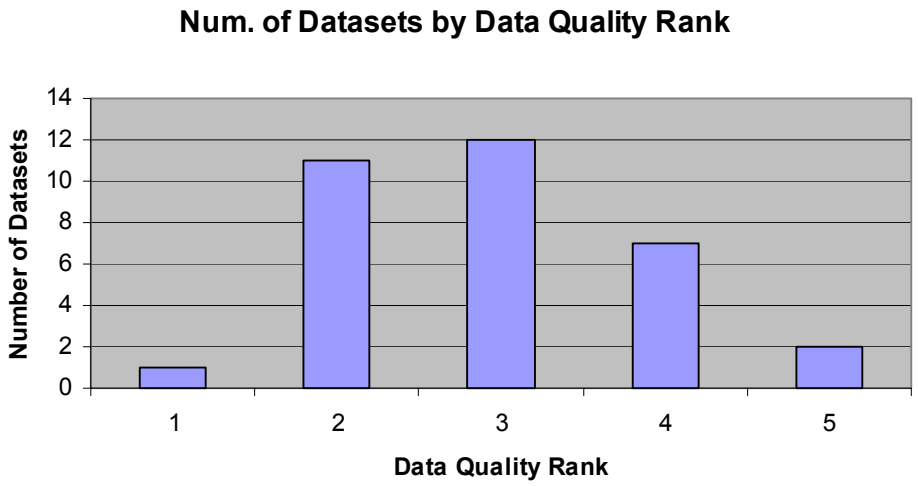

Fig. 2 Quality ranking of their datasets by councils.

Table 5 Criteria for ranking data quality measures in place for council datasets

Rank	Criteria
5	Robust measures are in place to ensure quality of data. Data have defined certainty statements and data collection methods are subject to continuous monitoring to ensure standards. Data undergo quality control procedures before being introduced into the data management system. There is a rigorous process for ensuring the data retain integrity (complete and absent from introduced errors). The quality assurance and control systems are externally accredited and independently audited.
4	There are very good measures to ensure quality of these data. The data in this system have defined certainty statements and some measures are in place to monitor collection methods. Standard quality control measures exist for entry of data and these are enforced. A process exists to ensure data integrity; however, it is not rigorously enforced or monitored. Quality assurance and control systems are documented but are not externally accredited or audited.
3	Reasonable measures are in place to ensure quality of data in the system. Data have defined certainty statements; however, techniques to monitor collection processes are relatively ad hoc. There are some basic quality control measures used when data are introduced to the system. There are no formal processes in place for ensuring data integrity; however, we believe the data do not suffer from significant introduced errors.
2	There are rudimentary quality control measures for this system. Processes for collection of data are relatively standardised and coordinated; however, there are no certainty statements in place for these data. No quality control measures exist for checking of data before these enter the system.
1	There are no quality management processes in place for these data. There is no measured level of certainty, and we are unable to guarantee that the same measures and instrument calibrations occurred for all the data in this system.

The datasets that ranked a 5 were the Greater Wellington Possum Residual Trap-catch dataset and the Environment Bay of Plenty Biodiversity database (NB this database has not yet been built). All but one of the datasets scoring a 1 or 2 were the small, generally spreadsheet-based datasets reported by Greater Wellington. Greater Wellington's larger datasets all scored a 3.

Datasets that were ranked a 4 were: Auckland Regional Council's Natural Heritage Database, Environment Bay of Plenty's Land Resources Database and Maritime Wetlands Database, Gisborne District Council's PMA Database, Marlborough District Council's QDAS, Otago Regional Council's Biomonitoring Database, and Taranaki Regional Council's Freshwater Biological Database.

It is important to note than in the surveyors' experience individual responses to this type of question can be quite variable. In particular those who appear to have a very good understanding of what data quality means are likely to rank their datasets lower than those who have less of an understanding.

Figure 3 shows quality rankings according to dataset type.

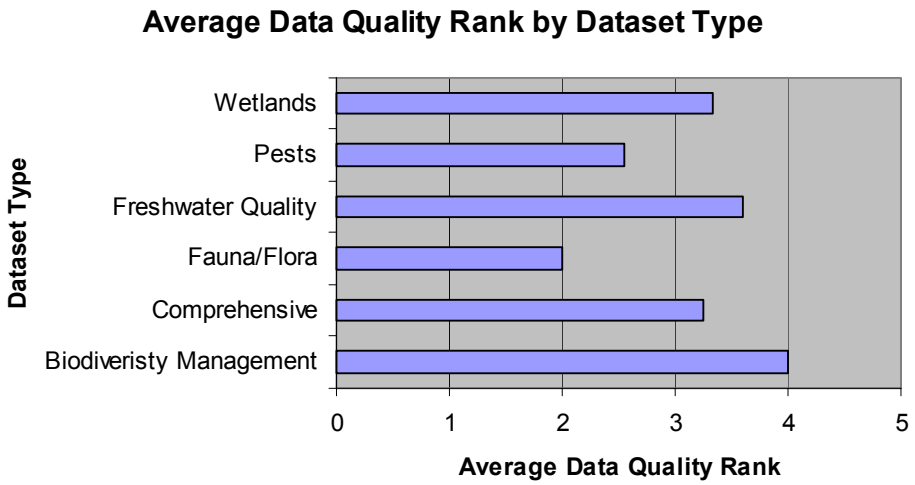

Fig. 3 Quality ranking by dataset type.

Biodiversity Management and Freshwater Quality datasets seemed to have the highest level of data quality measures. This is likely to be because these are generally larger datasets with quite particular resource management uses. Comprehensive datasets, while large, perhaps have their quality rankings dragged down slightly by the wider range of data stored.

In terms of organisations' average data quality ranks, no particular patterns emerged. Many of the smaller councils had datasets that they ranked as a 4. While this could possibly be due to the relative levels of understanding of data quality, this seems unlikely as an across-the-board generalisation. In the cases of Marlborough District Council's QDAS, Taranaki Regional Council's Freshwater Biological Database, and Environment Bay of Plenty's datasets for example, the associated comments they made gave credence to the veracity of their high data-quality rankings. It should be noted that those ranked as high quality by smaller councils tended to be predominantly Freshwater Quality datasets. Also, with the exception of Gisborne District Council, where the smaller councils listed only one dataset these were all Freshwater Quality ones.

4.4 Future development

A number of questions were asked in relation to future development plans. These included how often data were updated, what any potential risk of loss was, proposed development plans for existing databases, new databases planned, and what hardware and software platforms were being used.

Maintenance and future certainty

Respondents were asked to describe for each dataset the extent to which the data were continuously being maintained, and how certain the future of the data was, i.e. to what extent was it at risk of being lost?

There were no instances where data seemed at risk of being lost. It is possible of course that any datasets at risk of being lost were simply not known about or reported on. Almost all datasets in the survey were said to have a reasonable or high level of future certainty.

Twelve of the datasets were reported to be updated at least weekly, three were updated annually or 5-yearly, two were not updated, and two were only likely to be updated if the area was resurveyed (Environment Bay of Plenty's Maritime Wetlands Database, and Environment Canterbury's Wilding Conifer Database). Three were updated on an ad hoc basis (all significant natural area datasets that were updated when a new area was added or boundaries changed). For 10 of the 34 datasets

surveyed the respondent was uncertain how frequently they were updated. These were all smaller datasets listed by Greater Wellington.

Development path

Respondents were asked to explain the proposed development path for each of their databases. A number of databases were reported to be serving their purpose well, with no plans to develop them further. These included GDC PMA database, Horizons Natural Areas Database, MDC SIGAreas, MDC QDAS, NRC Sites, NRC Pest Eradication Contracts, and NRC Biosecurity.

Where further developments to existing databases were planned these fell loosely into three categories: scope expansion, quality and usability improvements, and external accessibility.

Comments for databases planning an expansion of scope include:

- **ARC Natural Heritage Database** – 'The natural heritage database is an ongoing project. We will be working to refine the existing database particularly for indigenous vegetation layers, extent of wetlands, and to further develop underlying attribute information. The database is currently very 'vegetation'-focused, we will be looking to develop better information on native invertebrates, skinks and geckos. We will also be including mapping of threatened plants locations on regional parks, analysis of ecosystem fragmentation, mapping of corridors and restoration networks linking with care groups and landowners working on biodiversity restoration.'
- **EBoP Freshwater and Maritime Wetlands databases** – 'These databases will be upgraded and included as a contributing database in the biodiversity databases and interfaces project.'
- **ECan Wilding Conifer Database** – '…intend to resurvey the whole region with the same method, create a new database, then can compare time 1 and time 2. Documenting control operations that have taken place as part of our regional pest management strategy so can look at the effect of controls.'
- **ORC Biomonitoring Database** – '…global dataset for the whole of council. Delivering hydrometric data on the web now, intend to make all our datasets available. In full review at the moment, June 2006 to begin to transfer data. May do mapping projects for each of our estuaries, and probably do a better job of wetlands in the medium term.'
- **GW KNE Geodatabase** – '…plans (still in their infancy) to expand to incorporate the database to capture much of the biosecurity biodata including casual observations.'

For some datasets plans were reported around improvement of data quality, management practices, and usability. These included:

- **EBoP Land Resources database** – '…ongoing review and upgrade as management processes mature. Add monitoring data for effectiveness measurement.'
- **GW Pest Plant datasets** – '…pest plants database to be upgraded to facilitate extraction of data through GIS GUI and address recently recognised data integrity shortfalls.'
- **WCRC Macroinveterbrates Spreadsheet** – '…no concrete plans, but would like to change the format in the future to make it more user-friendly for analysis and state of the environment reporting.'

For a number of datasets plans to move to server-level databases and make data externally accessible were reported. Comments included:

- **MDC Ecodata** – 'Ecodata are in development mode currently. Already mapped within science providers in Marlborough Region; we're running it internally here. Want to make it accessible over the Web.'

- **Horizons ecoBase** – '…it is a portable [Access] database currently. We plan to integrate it into the main council databases and make some of the data Web-accessible.'
- **TRC Freshwater Biological Database** – '…is currently on our openVMS, in Powerhouse, moving to SQL-Server in Powerbuilder on Windows.'

New biodiversity databases

Respondents were asked: *What new databases that will contain biodiversity-related data do you plan to develop in the future?*

Six organisations mentioned plans to develop new databases. These are as follows:

- **EBoP Biodiversity database** – '…a new database to record biodiversity information (wetlands, land, lakes, rivers and streams, marine, geothermal, estuaries and underground) for general to specific monitoring data. This will store data on values, extent, condition, management and monitoring. The new biodiversity databases will use a data-depreciation process to manage competing data, and metadata will be filled in for all data in the databases. Database administrator will ensure that all data are collected in accordance with standard methodologies.'
- **GW Coastal Habitats Database** – 'There is the potential for the need for a coastal habitats database. Creation of such a database will only come about if we cannot easily access or cannot rely on other coastal datasets.'
- **MDC Macroinvetebrate Pictures – '…**a small photographic database of macroinvetebrate pictures.'
- **NRC Wetland monitoring Database** – '…a new database based on WONI. This will involve lakes monitoring (biodiversity, biosecurity and water quality information) using NIWA standards (LakeSPI) and ANZECC water quality guidelines, and estuarine monitoring using Estuarine Monitoring Protocol (Cawthron Inst.).'
- **ORC Pest Species** – '…compliance officers currently maintain a dataset of pest species (plant and animal), e.g. maps of gorse spreading, this will become a full database in the fullness of time.'
- **TRC Pest Plants** – '…will be doing one with Pest Plants to record them geospatially – just started planning this.'

A general trend observable here is that some smaller councils are planning to expand from just water quality datasets into pest management and wetlands datasets.

Hardware and software platforms

Respondents were asked: *What hardware and software platforms are you using to store and manage your biodiversity-related data? To what degree do you consider these to be adequate for what you want to do in the future?*

The most common platforms were Microsoft SQL-Server, Access, and the ESRI Arc GIS tools (Table 6). All are 'proprietary' systems, which appear to be the de facto standard in regional councils. Other technologies that were mentioned include Inmagic, Powerbuilder, and Pathway/WorkSmart (GEAC).

Table 6 Technology platforms currently used for council databases.

Technology	Dataset
Microsoft Access	GW Wetlands Database, Horizons ecoBase, MDC Ecodata, MDC SIGAreas, MDC QDAS, ORC Biomonitoring Database
Microsoft SQL-Server	EBoP Biodiversity Database, Horizons Natural Areas Database (Powerbuilder front end), NRC Sites (Pathway/WorkSmart (GEAC)), TRC Freshwater Biological Database (Powerbuilder)
ESRI Arc GIS tools	ARC Natural Heritage Database (ARCView and Inmagic), ECan Wilding Conifer Database, GDC PMA Database, GW Wetlands Database, NRC Sites, ORC Biomonitoring Database

4.5 Integration and interoperability with major biodatabases

The topic of integration/interoperability with major biodatabases held by CRIs, Government departments and universities was raised. Questions were asked on: ability to exchange basic name, date and location data; other data fields of research interest; barriers to integration/interoperability; benefits of integration/interoperability; and use of external systems.

Ability to exchange basic minimum data

Respondents were asked the following question:

If you were able to provide data to, or integrate your databases with, major biodatabases held by central government and major research institutions, how would you rate your ability to provide or exchange very basic collection or observation data including at a minimum:

- *Name of the taxon to which the organism has been assigned*
- *Location at which the specimen was collected or the observation made*
- *Date on which the specimen was collected or the observation made*
- *Where the specimen is held and how to access more information.*

Some respondents said their datasets were already providing data to major biodatabases. These were as follows:

- **ARC Natural Heritage Database** – '…some of our data are already integrated with national databases. E.g. Our PNAP survey information, permanent vegetation plot information is fed into the NVS database, and freshwater fish and aquatic plant info into the NIWA databases, our wetland layer has been provide to DOC.'
- **MDC SIGAreas** – '…our fish data goes in to FBIS. I worry about misidentification, the level of expertise varies so much and I wonder that a lot of stuff that goes into FBIS is wrong. I could get the genus but would never attempt to get the species, just wouldn't have the expertise.'
- **TRC Freshwater Biological Database** – '…have sent quite a lot of data to various consultancies (including Cawthron) in the past for differing purposes. We retain all our specimens here.'

For nine datasets respondents said it would be relatively easy to provide data to national datasets. These included MDC SIGAreas, TRC Freshwater Biological Database, ARC Natural Heritage Database, EBoPlenty Biodiversity Database, ECan Wilding Conifer Database, Horizons Natural Areas Database, Horizons ecoBase, NRC Sites, ORC Biomonitoring Database. Most of these said they could 'copy data over' but not connect in real time. Horizons ecoBase was the only one that

mentioned ease of integration/interoperability due to their use of international data exchange standards.

Some cited difficulties or challenges around data management issues including:

- **GDC PMA database** – '…difficult because in lots of separate files.'
- **GW Wetlands Database** – '…integration will probably be only by way of 'cutting and pasting' into other datasets, e.g. pest plants database information was copied straight into Bioweb. We are aware of the challenges of concurrency and the risk of double counting the same entity, but have no formal policies in place to avoid this. This would be a reasonable benchmark for other datasets – i.e. we could do wholesale handover of data, but anything else would be tricky. Some data we are collecting goes directly into national databases.'

Other research-useful fields

For each type of primary collection or observation record respondents were asked what other fields they had that might be of benefit to major biodatabases (e.g. might be of use in research work done by scientists in CRIs and universities). This was a complicated question and most respondents did not have time to exhaustively list all fields. Many gave representative examples of other fields and these are listed in section 0 below.

Common responses included abundance, environmental conditions (e.g. temperature), location data, and threats. Most respondents listed between 5 and 20 fields.

Barriers to integration/interoperability

Respondents were asked: *What are the barriers you would perceive to integrating your data with major biodatabases? (e.g. legal or intellectual property constraints, data standards constraints, technical connectivity, etc.).*

Responses ranged across all of these types of barriers. Three organisations cited data exchange standard constraints. Some reported potential data quality issues, especially for older data. Four mentioned legal constraints, all related to data collected on private land. None reported intellectual property barriers. Three cited technical connectivity constraints. Four said that time and money would be a significant constraint, and that this sort of activity would be outside the remit of rates collected by councils. One smaller council said that extra funding wouldn't necessarily help as even with extra money it would be hard to get staff resources prioritised onto an activity like this. Some mentioned that most of the barriers became irrelevant in the case of systems like FBIS where they were able to enter data directly into that system rather than develop their own one.

There were no real correlations between the types of barriers mentioned and dataset type or organisation size.

Value of integration/interoperability

Respondents were asked what in their perception would be the value of integrating bio data with any of the major biodatabases. They were asked to respond in terms of value to their organisation, and value to other users of the data.

In almost all cases respondents were positive about this question and could list several potential benefits. Benefits were focused primarily on themselves and other end-users of the data. Several did mention benefits for researchers but none went into any detail on this.

Regional comparison of biodiversity data was the most common benefit with eight organisations specifically mentioning this. Two smaller councils said this might not be useful or would be difficult because of their unique environmental conditions or specialised monitoring methods. One said that it was important to be careful that the local knowledge surrounding the data was conveyed when data were shared outside the region, to prevent misinterpretation.

Comparison of biodiversity practices between regions was also cited as a potential benefit, i.e. learning what other people are doing. 'Getting more data on what's going on in our region' was a related comment. This was focused on situations where DOC (in particular) and CRIs (sometimes) had data that overlapped with theirs. One district council commented that they could see more benefit in integrating their data with DOC than with CRIs. Another district council was wary about using new national systems: 'If there was value, if everything was set up and ready to go, but don't want to double enter our stuff at all. Unless it's been going for a while we don't want to be on the cutting edge of this stuff. We're a small council with a tiny amount of ratepayers. Would follow after the big regional councils.'

Placing data in a national context, assessment of trends or biodiversity health nationally, and SoE monitoring were also mentioned several times. One respondent saw the potential for the eventual emergence of a 'meta-system': 'In the future I see our database being a node within a much larger system. We collect data at our scale, feed it up and access data at neighbouring regions. Biodiversity doesn't stop at our boundaries; we're a part of a much larger system.'

A small number perceived data quality and improved standards adoption as potential benefits: 'It also has implications for data quality. Integration encourages standards adoption, which in turn encourages better data collection practices in regional councils and other organisations such as NZERN.'

As an overall generalisation councils that had 'comprehensive' or biodiversity monitoring databases mentioned more benefits and tended to talk more about the 'national picture'.

Use of external systems

The following question was asked: *Do you currently store any data in any systems external to your organisation, in particular in the major biodatabases held by central government and major research institutions? For example NIWA's New Zealand Freshwater Fish Database (NZFFD) records the occurrence of fish in fresh waters of New Zealand, and is contributed to by 80 organisations external to NIWA including regional councils. If you do, what are these systems, and in general terms what data do you store in these?*

Almost all councils surveyed said they stored some data in external systems. Two small councils said they did not. It should be noted that this was by no means an exhaustive survey of external data storage and councils not mentioning particular systems does not conclusively mean they did not use those systems.

NZFFD/FBIS was the most common with nine councils (of all sizes) mentioning this (three explicitly referred to FBIS and/or the aquatic plant database). The three large councils said they all stored data in Landcare Research's National Vegetation Survey Databank (NVS). Others were mentioned just once including Landcare Research's New Zealand Arthropod Collection (NZAC), Fungal Herbaria, and Allen Herbarium. The following DOC datasets were mentioned once or twice each: threatened species database, weeds database, WONI (Wetlands of National Importance), Pest fish database.

From Table 7 it can be seen that currently small councils report storing only freshwater data in external systems. Large councils all reported storing terrestrial biodiversity data in external systems.

Interestingly only two report storing pest data in external systems, and these are both DOC systems.

Table 7 External database systems mentioned as used by councils.

External system	No. of mentions	Size of council mentioning
NIWA NZFFD/FBIS	9	3 large, 2 medium, 4 small
Landcare Research NVS	3	3 large
Landcare Research NZAC	1	1 large
Landcare Research Allen Herbarium	1	1 large
Landcare Research Fungal Herbaria	1	1 large
DOC threatened plants database	2	2 large
DOC weeds database	1	1 large
DOC WONI (Wetlands of National Importance)	1	1 medium
DOC Pest fish database	1	1 small

4.6 Integration and interoperability with other regional councils

The topic of integration/interoperability with other regional councils was raised. Questions were asked on benefits of integration/interoperability, value of sharing technology, and whether there had been any occurrences of reinvention of the wheel between regional councils.

Value of data integration/sharing

Respondents were asked: *If you were able to integrate data with, or access data from other regional councils would this be of benefit? What would be the value to you in doing this? What kind of data might you exchange?*

Many had already touched on this in responses to the question on benefits of integration/interoperability with major biodatabases. Nine (out of the 11 that responded) councils said there would be some value. Two (both district councils) said there would be little or no benefit. Three councils specifically mentioned benefits in terms of managing cross-boundary issues with neighbouring regions including pest management, biosecurity, biodiversity values, and shared ecological districts. Three others mentioned comparison benefits. While not specifically asked this question, two councils said they already shared information with other councils.

There was no real correlation (with the exception of the two district councils) between organisation size and whether or not respondents perceived benefits of data integration/sharing with other councils.

One respondent cited benefits in terms of data and resource management practices and said: 'There would be benefits in terms of the fact that we all deal with similar problems. Exchanging data would also lead to the exchange of knowledge about good data management practices, and also knowledge about resource management. Data are potentially a carrier wave for other knowledge. There is not yet enough cross communication between councils in terms of what can be done, and how to develop things further. It is improving but it is based on the people involved. There is no means of doing this consistently or nationally. At the moment leadership is very dependent on the person driving biodiversity management in the organisation. Good biodiversity data management isn't institutionalised yet. There is a high risk of loss of knowledge and capacity if people leave.'

Value of sharing technology

Survey respondents were asked what value they saw in sharing technology infrastructure with other regional councils (e.g. centralised servers, centralised system, decentralised copies of a system developed for several regional councils).

All of the respondents were highly negative about the prospect of sharing infrastructure with other regional councils. Some were supportive of the notion of centralised systems, some believed this would not cater to their unique regional needs. Some small and medium-sized councils could see the benefit in central systems and/or distributed copies of systems in terms of reduction of cost or improvement of data quality.

Other comments of interest were:

- 'Regional consortiums currently share access and purchase of orthoimagery. A centralised system would remove control and ownership over database management. The data are integral to the work we do and meets our own specific needs, e.g. mapping small fragments of vegetation and trees within a highly urban setting. Different to mapping and recording biodiversity in a big-sky South Island landscape.'
- 'Some districts within the region that presently don't have the technological/knowledge capacity would be able to better participate in regional decision-making. With other regions, it could avoid duplication.'
- 'We don't consider that this would have any value for Environment Bay of Plenty. However, we are willing to share database design with other regional councils if it is useful for them.'
- 'Of value in reducing cost, looked at this with Scott Crawford at Environment Southland. There are some challenges politically.'
- 'Peer-to-peer-style integration of data perhaps. Fairly independent systems linking up and exchanging data rather than a large centralised system.'

One conclusion that can be drawn from responses to this question is that there is no 'one size fits all' approach to biodiversity data. Centralised systems are likely to work well for some types of data, but perhaps not for others. Teasing out the difference between actual regional variation and a tendency to want to 'own the system' for its own sake is a difficult business. On the one hand organisations rave about how brilliant NZFFD is; on the other hand they are fearful about loss of control for other types of data.

'Reinvention of the wheel'

As the final question in the survey respondents were asked: *Do you have any examples of unintentional, or unavoidable (for whatever reason) reinvention of the wheel in terms of biodatabases between different regional councils? If so, please give a short summary of these.*

Six councils said they weren't aware of any reinvention of the wheel. Five of these were small councils and one said that they didn't know what systems other councils had.

Four councils said that they thought overlap or reinvention of the wheel happened more between regional councils and other agencies, in particular with DOC, MfE, and some Landcare Research systems. Specific comments about this included:

- **Gisborne District Council** – 'Seeing this sometimes in central government working with iwi groups – certainly an issue with land-use stuff. Often money spent on collecting information that we already hold. Not much consultation to see what's there already. Biodiversity condition fund was done independently of councils, has been some duplication there.'

- **Otago Regional** – '…more with major providers than other regional councils. We proposed building a complete Otago region fisheries database because we knew there was data in Otago Uni and DOC that wasn't in national databases. This resulted in those data getting into the national databases and us not having to build the database. Landcare Research land cover database and LENZ seem to duplicate components. Needs to be more direction to councils as to what they might use those databases for. OBI process may help but there are constraints in terms of councils being reluctant to support things that will favour one ratepayer over the other. Often the information on what you shouldn't use data for isn't provided in the end-products and that can lead to errors. It's easy getting the data in there, much harder getting the metadata/information/knowledge in there. As soon as you get spatial data in you get far more risk of misinterpretation being made.'
- **Environment Canterbury** – ' seems to be a certain amount of overlap with other agencies, e.g. something Horizons have done for wetland monitoring is similar to something MfE have done. Forest remnants horizons have done seems similar to something done by Peter Hanford in Kapiti Coast.'

Some comments were made about reinvention of the wheel between 'comprehensive' or biodiversity management type databases. These included:

- **Greater Wellington** – 'Horizons ecoBase and KNE are fairly similar, both were set up at roughly the same time. Horizons had thought about how to structure before collecting it; GW had collected it first. Birds data could be a duplication of effort as other people are sampling the same areas, and also duplication of databases and database design effort as there is no national birds database.'
- **Environment Bay of Plenty** – 'A few regional councils have developed biodiversity databases, e.g. Horizons. I am sure that we, in the development of our biodiversity database, had to address similar issues to them. We have both ended up with very different solutions. The other regional councils that are determining the best way to design a biodiversity database (e.g. ARC and EW) will address similar issues to us and undoubtedly will come up with different solutions. There is nothing wrong with different solutions for different regions. However, I do consider that more collaboration between regions would assist everybody come up with a solution that suits their region and reduces their workload.''

Taranaki Regional Council thought there might have been reinvention of the wheel with macroinvertebrate data. They said: 'We had a lot of queries about our macroinvertebrate database early on. Now it seems that others are developing their own. We think ours is unique and of high quality both in terms of the data, and of the database system itself.'

A number commented specifically on the need for more communication and sharing of knowledge about what others are doing:

- **Auckland Regional Council** – '…we share ecological survey information with other councils (e.g. Env Waikato, Northland DOC Conservancy and Auckland DOC Conservancy) and work in partnership with TAs in our region to undertake survey and monitoring work – there has not been any duplication of effort on this. There is merit, however, in sharing database set-up methodologies and creating common databases (e.g. wetland extent, vegetation extent).'
- **Environment Canterbury** – '…would like to see a list of the databases that different organisations have – so if you were interested in one you could find one.'
- **West Coast Regional Council** – '…no reinvention that I'm aware of in terms of systems. Certainly true in terms of approaches and techniques others are using and innovation in these. Only find out by going to conferences.'

5 Appendix

This appendix contains references used in the preparation of the report, a set of statistical data on regional council funding, and detailed findings from this survey.

Detailed findings include a full list of datasets by dataset type and sorted by organisation size, and a list of research useful fields that could be provided to major biodatabases. Full information collected for each council and dataset including metadata and responses to survey questions are also provided.

5.1 References

New Zealand Biodiversity Strategy – http://www.biodiversity.govt.nz/

TFBIS – http://www.biodiversity.govt.nz/land/nzbs/tfbis/tfbis/index.html

GBIF – http://www.gbif.org

New Zealand GBIF node – http://www.gbif.org.nz

Darwin Core – http://darwincore.calacademy.org/

ABCD standard – http://bgbm3.bgbm.fu-berlin.de/TDWG/CODATA/Schema/default.htm

Other TDWG standards – http://www.tdwg.org/subgrops.html

New Zealand Geospatial Metadata standard (NZGMS) –

http://www.e-government.govt.nz/interoperability/nzgms.asp

New Zealand Government Locator Service (NZGLS) – http://www.e-government.govt.nz/nzgls/

5.2 Glossary

Research data – defined as in the U.S. National Institutes of Health definition of final research data: 'the recorded factual material commonly accepted in the scientific community as necessary to validate research findings'. It should be noted that in this paper the term 'data' is used in the scientific, plural sense, i.e. this 'datum', these 'data'.

Biodata – research data relating to biology and/or biodiversity, often relating to collections or observations of plants, mammals and invertebrates.

Biodatabase – a database containing biodata.

Geographic Information System (GIS) – a computerised system for combining, displaying, and analyzing geographic data. GIS produces maps for environmental planning and management by integrating physical and biological information (soils, vegetation, hydrology, living resources, and so forth) and cultural information (population, political boundaries, roads, bank and shoreline development, etc.).

Collections – a collection is a set of specimens collected in the field and held in a particular institution. For some collections, individual collection items are 'lots' which contain more than one specimen. Each collection item is typically labelled in some way. In well-managed collections collection items are independently identified, vouchered and metadata stored in a collection management database. Collections are often held in climate-controlled environments. For this report a dataset can be a collection dataset (i.e. it contains records of many individual collection items). In this report phrases like 'number of collections' is shorthand for number of collection items, i.e. five collections is synonymous with five specimens.

Observations – an observation is a record of a sighting of a particular organism in the field. In this report some datasets are referred to as observation datasets (i.e. containing records of many observations).

Taxonomic names – the scientific name of a plant or organism, including its place within the Linnean hierarchy (Kingdom – Phylum – Class – Order – Family – Genus – Species), and often including vernacular names and synonyms.

Taxa – (singular = taxon) the named classification unit to which individuals or sets of species are assigned, such as species, genus and order.

Darwin Core – a simple set of data element definitions designed to support the sharing and integration of primary biodiversity data.

ABCD – the Access to Biological Collections Data (ABCD) Schema is an evolving comprehensive standard for the access to and exchange of data about specimens and observations (a.k.a. primary biodiversity data).

Global Biodiversity Information Facility (GBIF) – GBIF is a distributed query network comprising data provider nodes all around the world. It exists to make the world's biodiversity data freely and universally available by developing biodiversity informatics tools to provide Web access to primary information on the world's organisms.

Data dictionary – a controlled list of items that restrict entry of data into a particular field in a database to only the items in that list.

Informatics – A field of study that focuses on the use of technology for improving access to and utilisation of information. Health informatics, for example, is the systematic study of information in the healthcare delivery system, how it is captured, retrieved, and used in making decisions, as well as the tools and methods used to manage this information and support decisions.

Middleware – Software that sits between two or more types of software and translates information between them. Middleware can cover a broad spectrum of software and generally sits between an application and an operating system, a network operating system, or a database management system. In the context of research data middleware applications often include tools that enable indexing, archiving, discovery, analysis, integration, management and preservation of large heterogeneous distributed data repositories.

Web services – Web services let computers talk to one another over the Internet, allowing computer programs to exchange information by eliminating barriers such as different hardware platforms, software languages, and operating systems that usually make different programs incompatible. Web services make it easier to share information, data, and services, as well as making it cheaper and easier for organisations to work with on-line partners.

5.3 MoRST statistical data on regional councils

These data were collected during MoRST's recent evaluation of the environmental output class. It should be noted that figures in relation to research and investigations (R&I) spend, and scientists employed is fairly subjective.

Council	Approximate annual expenditure on R&I and total expenditure	Approximate % of R&I eontracted	Scientists employed
Regional councils	Note: at the time of publication these figures were still being checked to ensure that the definitions were evenly applied.		
Auckland	$2.5 million (varies between $1.8 and $3 million), $175 million	90%	21
Bay of Plenty	$1.7 million (about to increase for lake research), $48 million	60%	6
Canterbury	$600,000, $64 million	50%	11
Hawke's Bay	$500,000, $30 million	10%	6
Manawatu-Wanganui	$280,000, $28 million, of which $6 million on environment	40%	7
Northland	$200,000, 12 million, $5.5 million on environment.	50%	6
Otago	$150,000 (has been higher recently due to Grow Otago), $14 million	80%	6
Southland	$500,000, $18.5 million	10%	4
Taranaki	$250,000, $10 million	35%	10 (20% on R&I)
Waikato	$3 million, $52 million. Resource Investigations spends $5.8 million; includes R&I and SoE.	40%	14
Wellington	$200,000, $114 million, including bulk water. Environment is $10.6 million	65%	5
West Coast	$80,000, $6 million	20%	1
Unitary authorities			
Gisborne	$30,000 ($15–20,000 on coastal hazards). $48 million, $4.5–5 million on regional function	90%	0
Marlborough	$1 million (has gone up by 30% over last few years), $40 million	80%	5 (75% on R&I)
Nelson	$180,000, $45 million. Smallest unitary so district function predominant	70%	0
Tasman	$730,000, $46.5 million	30%	4.5 (50% on R&I)
City councils			
Christchurch	About $1 million, $288 million	Most	1
North Shore	$300,000, $167 million	30%	0
Waitakere	$325,000, $150 million	100%	0
Wellington	c. $250,000 – hard to assess as spread across Council, $260 million	Not sure	Not applicable
District councils			
Hastings	$30–40,000, $55.5 million	60%	Not applicable
Wanganui	$10–20,000, $30 million	100%	Not applicable

5.4 Detailed findings

The following is a record of survey response data for each individual organisation and dataset.

Dataset list

Type of dataset	Organisation and dataset title
Comprehensive	Auckland Regional Council Natural Heritage Database Greater Wellington Key Native Ecosystems (KNE) Geodatabase Horizons ecoBase Northland Regional Council Sites
Biodiversity Management	Environment Bay of Plenty Land Resources Database Environment Bay of Plenty Biodiversity Database Gisborne District Council PMADdatabase Horizons Natural Areas Database Marlborough District Council SIGAreas
Wetlands	Environment Bay of Plenty Freshwater Wetlands Database Environment Bay of Plenty Maritime Wetlands Database Greater Wellington Greater Wellington Wetlands Database
Freshwater Quality	Greater Wellington Fish and Macro-invertebrates Marlborough District Council QDAS Otago Regional Council Biomonitoring Database Taranaki Regional Council Freshwater Biological Database West Coast Regional Council Macroinveterbrates Spreadsheet
Pests	Environment Canterbury Wilding Conifer Database Greater Wellington Pest Plant extent in regional parks (Parks & Forests) Greater Wellington Pest Plant location and extent Geodatabase (Biosecurity) Greater Wellington Possum Residual Trap-catch data Greater Wellington Parks & Forests rat/rodent and mustelid data Greater Wellington Biosecurity rat/rodent data Greater Wellington Parks & Forests goat culling records Greater Wellington Goat spatial extent GIS layer (Biosecurity) Greater Wellington Various pest animal records not identified above Northland Regional Council Biosecurity Northland Regional Council Pest Eradication Contracts
Fauna/Flora	Greater Wellington Parks and Forests Bird Count data Greater Wellington Biosecurity Bird Count data Greater Wellington Various special (e.g. rata and rare/endangered) plant species records / Casual observations of rare/endangered animalia (including invertebrates) Greater Wellington Phenology data Greater Wellington KNE Invertebrate Monitoring Marlborough District Council Ecodata

5.4.2 Other research-useful fields

This is a full list of the fields given as examples from datasets surveyed that may be of interest to researchers in CRIs.

- **Auckland Regional Council Natural Heritage Database** – Threats (weed and pest information), Ecological significance, Vegetation types, Analysis of fragmentation
- **Environment Canterbury Wilding Conifer Database** – Species composition, age, density, if a plantation we record risk of spread from this plantation (with respect to hills, surrounding landuse, wind directions, age of trees, whether there has been fringe spread)
- **Gisborne District Council PMA database** – Species that was there, why it was important, relative abundance (in general terms)
- **Greater Wellington Wetlands Database** – Wetlands: Vegetation (structure, species from most to least dominant – top 5), assessment of its quality (incl. count of no of culverts etc.). Largely, our additional fields such as pest control history for certain areas of special habitat are outside the scope of TFBIS?
- **Horizons Natural Areas Database** – Management Info: info on the whole site: Owner, owner attitude to retirement and pest control owner actions, regional council actions, fence status, grazing status, pests and threat, legal protection and type. All the fields that get recorded for FBI plots; plot info: slope, grid ref, landform, vegetation, altitude, pest species (weeds and animals). And the tree info: Species, dbh, distance and direction from plot centre, stem/tree, commonness, canopy/understorey/emergent/isolated tree, dieback top/whole, browse top/whole, canopy density, stem use, flowering & fruiting. FBI viewpoint plots get a subset of the FBI plot info, but for more individuals. Understorey plots: recording species and height within 49-cm radius of peg, up to 1.35 m, within five height classes.
- **Horizons ecoBase** – Habitat types (50), structure (how high, levels of species), estimates of abundance (counts or groups or % depending on size of habitats), threats (animal, human) – all part of rapid ecological assessment. For plot data, all from forest monitoring toolkit for transects etc.
- **Marlborough District Council SIGAreas** – Abundance, significant areas – level of rareness, how robust the site is (ranking), Geoff Walls – one of the better ecologists – does this for this. Representativeness, rarity, diversity and pattern, distinctiveness, size & shape, connectivity, sustainability. Each ranked individually, then an overall significant rank of the site.
- **Northland Regional Council Sites** – For the current dataset only basic density/population data but will include ecological monitoring data based on a national system developed by MfE (FORMAK). For future biodatabases we will collect data that follow national protocols and therefore will be of benefit to national databases
- **Otago Regional Council Biomonitoring database** – MCI, total richness, SQMCI (quantitive assessment), EPT, some with total numbers, water quality data (bi-monthly, so temporarily more detailed). Chris did a full stream database at Otago University. Most university data are never really published as it is done for specific outputs (student thesis).
- **Taranaki Regional Council Freshwater Biological Database** – MCI, how many found (relative abundance), sampling method, environment it was found in (water temperature, substrate type, abundance of periphyton, riparian vegetation cover), site (lat long location, height above sea level, distance to mouth of river), hydrological information (river flows etc.). Chemical data can be provided through link by site name to our chemical database.
- **West Coast Regional Council Macroinveterbrates Spreadsheet** – dissolved oxygen ph, tebidity, resuspendible solids, other physio chemical data.

(1) Auckland Regional Council

Natural Heritage Database

Title	Natural Heritage Database
Abstract	A comprehensive database of information on significant natural areas (collected from PNAP surveys and SSWI wildlife surveys), wetlands, geopreservation sites, layer of extent of existing indigenous vegetation, bibliographic database, restoration sites. It includes: a. data collected on significant natural areas by PNAP (Protected Natural Areas Programme) surveys (basis for schedules in regional and district plans, priorities for management, and planning for Regional Growth Forum) – PNAP survey data are also provided to NIVS; b. SSWI (Sites of Special Wildlife Interest) data; c. all known wetlands in the region (mapped from field surveys and/or aerial photography); d. geopreservation sites (sites of geological and landform significance); e. regional extent of indigenous vegetation (developed from a combination of data collected from ecological surveys, aerial photography); f. priority areas for possum control in the region (based on significant natural area surveys); g. bibliographic database – polygons are linked to a comprehensive bibliography of natural heritage information in the region h. mapping revegetation sites on Regional Parks (dates and species planted) We also have water quality, stream monitoring and freshwater invertebrate information collected by our Environmental Research team – held mainly as reports and spreadsheets (John Maxted has more information on this)
Purpose	Used on a day-to-day basis to identify significant areas, develop schedules in regional and district plans, basis for planning under Regional Growth Forum, assess environmental effects of resource consents, develop priorities for management effort (e.g. priorities for possum control, regional park management), provide information on natural areas to landowners and community groups, answer public enquiries.
Contact Person	Sam Hill, Natural Heritage Section, ARC
Organisation	Auckland Regional Council
Temporal Extent	Over 25 years of data
Spatial Extent	Auckland Region. DOC Auckland Conservancy has also used layers to provide Conservancy extent.
Number of Records	Over 6000 records
Approach to Metadata Management	Natural Heritage database is managed by the Natural Heritage team. The database is over 25 years old. It began as a bibliographic database of information on natural heritage in the region and as field survey information from PNAP surveys. It is now stored as polygons and fields in ARCView, linking to a bibliographic database in 'InMagic'.
Data Standards	**Rank:** 4 **Standards Used:** The Natural Heritage Database uses recognised national data management standards such as PNAP methodology, SSWI

	methodology and GRID Wetlands Classification **Notes:**
Data Dictionaries	**Dictionaries used:** Landcare Research Plant Names Database, LENZ, GRID wetlands methodology **Notes:**
Data Management Policy / Procedures	ARC has protocols over management of GIS datalayers and access to information. ARC Information Management is currently being reviewed.
Data Quality	**Rank:** 4 **Notes:** We have very good measures to ensure quality of data, however, there are limitations with accuracy of mapping, information collected at different times over many years, lack of standards for common vegetation classifications between ecological districts. There are gaps in our data, e.g. mapping vegetation extent, information for some ecological districts is older than others, mapping of all wetlands in the region.
Future Development	**Proposed development path for the database:** The natural heritage database is an ongoing project. **Extent the data are continuously being maintained:** The data are being continuously maintained and added to and the future certainty of the database is high. **Certainty of the future of the data:** There is no risk of it being lost and neglected unless Heritage Dept resources were cut significantly.

Development, Integration and Interoperability

The following are responses by Auckland Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	We will be working to refine the existing database particularly further refine indigenous vegetation layers, and extent of wetlands and further develop underlying attribute information. The database is very 'vegetation' focused – we will be looking to develop better information on native invertebrates, skinks and geckos Mapping of threatened plants locations on regional parks Analysis of fragmentation Mapping of corridors – restoration networks – linking with care groups and landowners working on biodiversity restoration
Hardware and Software Platforms	ARCView and Inmagic
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Could be done through: Location of significant natural areas (Polygons), vegetation type information, species location, ecosystem type (e.g. wetlands) Some of our data are already integrated with national databases. E.g. Our

	PNAP survey information, permanent vegetation plot information is fed into the NIVS database, and freshwater fish and aquatic plant info into the NIWA databases, our wetland layer has been provide to DOC. **Other fields that might be of benefit to major biodatabases:** Threats – weed and pest information Ecological significance Vegetation types Analysis of fragmentation
Barriers to Integration with Major Biodatabases	There would be data standard constraints (e.g. vegetation classification systems) – while information has been collected using standard national methods – information has been collected specifically for each ecological district.
Benefits of Integration with Major Biodatabases	Biodiversity research, predicted species distribution data – vegetation layers, placing data in national context
Externally Stored Data	**Data currently stored in external organisation's systems:** NIWA freshwater fish database, NIWA aquatic plant database – All our freshwater fish and aquatic plant data collected by our Environmental Research team (Grant Barnes is the contact) is entered directly into the national databases maintained by NIWA NVS national vegetation database – Permanent vegetation plot data for Hunua Ranges, Tawharanui Regional Park – this is fed directly into NVS
Integration with other Regional Councils	**Benefits of data integration:** Would be of benefit for cross boundary issues, e.g. shared ecological districts, TA boundaries that cross regional boundaries We already share ecological survey information with other regional councils and DOC conservancies. Our databases, however, have been developed to meet our specific needs and uses in the region. The uses of the data are primarily within our region. We are the ones in contact with landowners and public about the biodiversity of our region. **Benefits of sharing systems/infrastructure:** Regional consortiums currently share access and purchase of orthoimagery. A centralised system would remove control and ownership over database management. The data are integral to the work we do and meet our own specific needs, e.g. mapping small fragments of vegetation and trees within a highly urban setting. Different to mapping and recording biodiversity in a big-sky South Island landscape. **Examples of unintentional or unavoidable reinvention of the wheel:** No – we share ecological survey information with other councils (e.g. Env Waikato, Northland DOC Conservancy and Auckland DOC Conservancy) and work in partnership with TAs in our region to undertake survey and monitoring work – there has not been any duplication of effort on this. There is merit, however, in sharing database set-up methodologies and creating common databases (e.g. wetland extent, vegetation extent)

(2) Environment Bay of Plenty

Freshwater Wetlands Database

Title	Freshwater Wetlands Database
Abstract	Data held of freshwater wetlands, collected at different times for different purposes including from SNA and PNA reports and field visits
Purpose	To hold existing data on freshwater palustrine wetlands
Contact Person	Kate McNut
Organisation	Environment Bay of Plenty
Temporal Extent	1972 to current
Spatial Extent	The Bay of Plenty Region
Number of Records	341 freshwater wetlands each with limited data on description, hydrology/geology, flora and fauna, water quality, land use, values and threats
Data Standards	**Rank:** 2 **Standards Used:** Internal **Notes:**
Data Dictionaries	**Dictionaries used:** **Notes:**
Data Management Policy / Procedures	
Data Quality	**Rank:** 3 **Notes:** But during process of review and joining the biodiversity database project this will move to a ranking of 5
Future Development	**Proposed development path for the database:** The database will be upgraded and included as a contributing database in the biodiversity databases and interfaces project. **Extent the data are continuously being maintained:** Data not regularly updated. Once this database is part of the biodiversity database module more regular updating and maintenance of the data will occur. **Certainty of the future of the data:** Very low risk of being lost

Maritime Wetlands Database

Title	Maritime Wetlands Database
Abstract	Records vegetation mapping from combination of permanent plots, photo survey and field visits.
Purpose	Monitoring the extent and condition of maritime wetlands (maritime wetlands are those that close or contiguous with estuaries)
Contact Person	Stephen Park
Organisation	Environment Bay of Plenty
Temporal Extent	2000 to current (although includes an estimate of 1840 wetlands vegetation)

Spatial Extent	Bay of Plenty Region
Number of Records	About 2,500+ entries
Data Standards	**Rank:** 2 **Standards Used:** Internal **Notes:**
Data Dictionaries	**Dictionaries used:** Landcare Research Plant Names Database **Notes:**
Data Management Policy / Procedures	
Data Quality	**Rank:** 4 **Notes:** But during process of review and joining the biodiversity database project this will move to a ranking of 5
Future Development	**Proposed development path for the database:** The database will be upgraded and included as a contributing database in the biodiversity databases and interfaces project. **Extent the data are continuously being maintained:** Resurvey will be undertaken **Certainty of the future of the data:** Very low risk of being lost

Land Resources Database

Title	Land Resources Database
Abstract	Records pest information (distribution etc.) and formal protection areas
Purpose	To assist protection of voluntary managed areas of biodiversity on private land
Contact Person	Wayne Smith
Organisation	Environment Bay of Plenty
Temporal Extent	Formal protection 20–30 years. Pests 4 years
Spatial Extent	Bay of Plenty Region
Number of Records	Approx 700 formal protection properties. Thousands of pest plant records.
Data Standards	**Rank:** 2 **Standards Used:** Internal **Notes:**
Data Dictionaries	**Dictionaries used:** Stream names (New Zealand Standard), Plant pest names from National Plant Pest Accord **Notes:**
Data Management Policy / Procedures	
Data Quality	**Rank:** 4

	Notes:
Future Development	**Proposed development path for the database:** Ongoing review and upgrade as management processes mature. Add monitoring data for effectiveness measurement. **Extent the data are continuously being maintained:** Regular update of data on daily/weekly basis **Certainty of the future of the data:** Very low risk of being lost

Biodiversity Database

Title	Biodiversity Database
Abstract	Records biodiversity information (wetlands, land, lakes, rivers and streams, marine, geothermal, estuaries and underground) for general to specific monitoring data. Stores data on values, extent, condition, management and monitoring
Purpose	To store all biodiversity data in one location (including monitoring data)
Contact Person	Current project manager – Sue Mavor
Organisation	Environment Bay of Plenty
Temporal Extent	Mid-1980s to current
Spatial Extent	Bay of Plenty Region
Number of Records	Not built or populated yet so don't know
Data Standards	**Rank:** **Standards Used:** **Notes:**
Data Dictionaries	**Dictionaries used:** Landcare Research Plant Names Database, Valuation New Zealand **Notes:**
Data Management Policy / Procedures	The new Biodiversity databases will use a data depreciation process to manage competing data and metadata will be filled in for all data in the databases. Data base administrator will ensure that all data are collected in accordance with standard methodologies.
Data Quality	**Rank:** 5 **Notes:**
Future Development	**Proposed development path for the database:** **Extent the data are continuously being maintained:** **Certainty of the future of the data:**

Development, Integration and Interoperability

The following are responses by Environment Bay of Plenty to questions on future biodiversity database developments and integration/interoperability issues.

Approach to Metadata Management	No authoritative schema but will have for the new Biodiversity databases
New Databases	Biodiversity database and interface as outlined earlier

Planned	
Hardware and Software Platforms	Currently we use a multiple layer of servers that have multiple level of backup. All running on Windows 2000/2003 software platform. Environment Bay of Plenty will continue to use this set up in the future.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Good **Other fields that might be of benefit to major biodatabases:** Best to come and see the databases you are interested in. Time does not permit us to list all the fields.
Barriers to Integration with Major Biodatabases	Privacy – Often data are collected under a partnership agreement with landowners (I.e. not to be shared with other organisations). Accuracy – some data, particularly older data has questionable accuracy.
Benefits of Integration with Major Biodatabases	Others would receive valuable regional data . We would receive some useful data.
Externally Stored Data	**Data currently stored in external organisation's systems:** NIWA's New Zealand Freshwater Fish Database, Pest fish database shared with Bay of Plenty Conservancy.
Integration with other Regional Councils	**Benefits of data integration:** Useful to get data on boundaries of the region. Pest data and biodiversity values would be useful. **Benefits of sharing systems/infrastructure:** We don't consider that this would have any value for Environment Bay of Plenty. However we are willing to share database design with other regional councils if it is useful for them. **Examples of unintentional or unavoidable reinvention of the wheel:** A few regional councils have developed biodiversity databases, e.g. Horizons. I am sure that we, in the development of our biodiversity database, had to address similar issues to them. We have both ended up with very different solutions. The other regional councils that are determining the best way to design a biodiversity database (e.g. ARC and EW) will to address similar issues to us and undoubtedly with come up with different solutions. There is nothing wrong with different solutions for different regions. However I do consider that more collaboration between regions would assist everybody come up with a solution that suits their region and reduces their workload.

(3) Environment Canterbury

Wilding Conifer Database

Title	Wilding Conifer Database
Abstract	Results of a survey of canterbury high country of distribution and density of wilding pine trees and possible plantation sources.
Purpose	To help direct control operations for wilding trees, prioritise sites for control, provide information to individual landowners (including DOC).

Contact Person	Philip Grove
Organisation	Environment Canterbury
Temporal Extent	One-off survey started in 1998 and finished in 2003.
Spatial Extent	2 million hectares of Canterbury High Country (from foothills east to Main Divide).
Number of Records	Approximately 1000 shape files with attribute tables, and lots of points.
Approach to Metadata Management	Technical Report filed in TRIM records management system.
Data Standards	**Rank:** 2 **Standards Used:** **Notes:** Weren't aware of any surveys like this. DOC have used it in the bigger picture analysis of trends and for local operational planning. DOC have been given the dataset as a whole.
Data Dictionaries	**Dictionaries used:** Tree species of half a dozen tree species (e.g. Douglas-fir, *Pinus*), Surrounding landuse intensive grazing, low-intensity grazing, no grazing), Spread risks on plantation (point scale). **Notes:** Internal only
Data Management Policy / Procedures	Think there is. Have a GIS specialist.
Data Quality	**Rank:** 3 **Notes:** 3–4 For field survey have a procedure manual with methodology and explicit limitations of the survey. Disclaimer – it might have changed since recorded. Entry of GIS data is done by a specialist under guidelines from the procedure manual.
Future Development	**Proposed development path for the database:** Intend to resurvey the whole region with the same method, create a new database, then can compare time 1 and time 2. Documenting control operations that have taken place as part of our regional pest management strategy so can look at the effect of controls. **Extent the data are continuously being maintained:** It's not, survey is finished, hopefully we'll do another. **Certainty of the future of the data:** Reasonable level of certainty.

Development, Integration and Interoperability

The following are responses by Environment Canterbury to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	Others at ECan – databases of wetlands, significant natural areas (a lot of it is data from DOC or its predecessors), own ones – records of noxious pests, bird habitats, acquatic plants, freshwater things (terrestrial has only really come in over the last 2 years).
Hardware and Software	ARC Info GIS database

Platforms	
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Only 6 species, all there in the files. Only identification of actual species on plantations. Only identified at family level for wilding observations (but within 6 or 7 species). **Other fields that might be of benefit to major biodatabases:** Species composition, age, density, if a plantation we record risk of spread from this plantation (with respect to hills, surrounding landuse, wind directions, age of trees, whether there has been fringe spread)
Barriers to Integration with Major Biodatabases	Would, in good conscience, need to get consent from landowners before making data publicly available and some may decline. Time and money.
Benefits of Integration with Major Biodatabases	For ECan – help us in choosing where to prioritise control operations (if we knew that there were threatened species, threatened by spread of trees), and nationally in terms of overall trends or control programmes in other regions. For others – as above, perhaps scientists (but maybe limitation in terms of the collection methods).
Externally Stored Data	**Data currently stored in external organisation's systems:** NIWA – NZFFD, Landcare NVS, DOC – rare plants, weeds etc.
Integration with other Regional Councils	**Benefits of data integration:** Some, but relatively minor **Benefits of sharing systems/infrastructure:** **Examples of unintentional or unavoidable reinvention of the wheel:** Seems to be a certain amount of overlap with other agencies. E.g. something Horizons have done for wetland monitoring is similar to something MfE have done. Forest remnants, what Horizons have done seems similar to something done by Peter Hanford in Kapiti Coast. Would like to see a list of the databases that different organisations have – so if you were interested in one you could find one.

(4) Environment Waikato

Environment Waikato did not provide a full response; however, the following are notes from email dialogue and reference documents sent by Karen Denyer.

List of biodiversity related databases taken from a report entitled: 'Identification of Significant Natural Areas in the Waikato Region using remote sensing and existing databases', S.M. Beadel, W.B. Shaw, October 2000:

- Taupo District Council Natural Areas Register
- Geothermal Vegetation of the Waikato Region
- Natural Heritage of Rotorua District
- Waikato Region Indigenous Vegetation (1840 and 1992)
- Consents Database
- Resource Consent Application Database
- Waikato Regional Coastal Plan

- Waikato District Native Vegetation Inventory

Comments from Karen Denyer:

'This is only up to 2000. We have since developed an inventory of geothermally influenced vegetation, and sites that are known or likely to meet our significance criteria in Otorohanga District. We also regularly update the QEII GIS layer.'

(5) Gisborne District Council

PMA database

Title	PMA database
Abstract	Protection Management Area database – areas are on the GIS system, ARCINFO (in district plan maps) as polygon boundaries, PMA (protection management area) name, and a number linking to a text file. The Textual records are stored on electronic file. It contains information from PNAP Protected natural areas surveys, plus some additional info we paid for ourselves. We have sometimes altered boundaries because they were wrong or have changed. It has descriptive information about significant natural areas, done by ourselves, DOC, Landcare Research. It includes vegetation types, where in the area they are, and their significance in the context of other areas in the vicinity.
Purpose	
Contact Person	Trevor Freeman
Organisation	Gisborne District Council
Temporal Extent	1980s to 2001
Spatial Extent	Gisbourne district
Number of Records	Not sure
Approach to Metadata Management	
Data Standards	**Rank:** 4 **Standards Used:** PNAP **Notes:**
Data Dictionaries	**Dictionaries used:** Landcare Names database, What's in normal topo sheets **Notes:** Because Landcare Research did a lot of the survey work
Data Management Policy / Procedures	Wouldn't often get changed. Where errors are identified these get changed along with DOC. No formal policies.
Data Quality	**Rank:** 4 **Notes:** Had to be done to certain standards, lots of quality control on the GIS system.
Future Development	**Proposed development path for the database:** None **Extent the data are continuously being maintained:**

	Unlikely to do further reviews/additions in the short/medium term (unless for example DOC reviewed some of their older PNAP surveys). **Certainty of the future of the data:** No risk of loss

Development, Integration and Interoperability

The following are responses by Gisborne District Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	Nothing
Hardware and Software Platforms	ARCINFO and Word docs, Excel files.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Difficult because in lots of separate files **Other fields that might be of benefit to major biodatabases:** Species that was there, why it was important, relative abundance (in general terms)
Barriers to Integration with Major Biodatabases	Technology barriers.
Benefits of Integration with Major Biodatabases	LENZ is too broad (and disparaging comments about this from DOC). More of interest to DOC. When we extended the PNAP surveys that was of interest to DOC. Possibly Fish & Game. If it went into a bigger system it would probably be DOCs.
Externally Stored Data	**Data currently stored in external organisation's systems:** None.
Integration with other Regional Councils	**Benefits of data integration:** Can't see benefit to us. Biodiversity has much less cross boundary issues that pest management etc. **Benefits of sharing systems/infrastructure:** Would only look at if there was a cost saving. **Examples of unintentional or unavoidable reinvention of the wheel:** None. Seeing this sometimes in central government working with iwi groups – certainly an issue with land use stuff. Often money spent on collecting information that we already hold. Not much consultation to see what's there already. Biodiversity condition fund was done independently of councils, has been some duplication here. Small, unitary authority, resources are short, don't have a scientific base at all, don't carry out our own research. Don't carry out any data collection in the biodiversity area. Only done under contracts, and normally driven by statutory requirements (district plan).

(6) Greater Wellington

Greater Wellington Wetlands Database

Title	Greater Wellington Wetlands Database
Abstract	Geodatabase (GIS data-entry and reporting, with link to Access database to hold non-GIS useable attribute data). For all known natural wetlands, and human-made wetlands of biodiversity value. Wetlands classified using the SMF classification system. Wetlands mapped, showing hydrological and vegetative extent of each wetland class, to delineate spatial extent of wetland habitats. Data include main vegetation, pest plant and animal infestation, human intervention (e.g. drainage, fencing), management needs. The database is in the latter stages of development.
Purpose	To provide data for State of the Environment monitoring / analysis of policy effectiveness, and to determine priorities for management.
Contact Person	Melanie Dixon (Policy Advisor (Wetlands Ecologist)), John Gibson (GIS specialist)
Organisation	Greater Wellington
Temporal Extent	2002–2004 snapshot with periodical updates for newly discovered wetlands. An intention to do 'snapshots' every 5 years.
Spatial Extent	Wellington Region.
Number of Records	Approximately 200 wetlands, with attribute data in 28 fields for each.
Data Standards	**Rank:** 4 with regard to data collection standards, 2 in terms of internal storage **Standards Used:** SMF Classification (original version) **Notes:** There has been a recent version change in the SMF classification, but it does not affect the integrity of the data collected.
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	Manual on data entry. GW Doc 220829. Fully complied with (only one data entry operator used)
Data Quality	**Rank:** 3 **Notes:** SMF Classification (original version) and training of staff. Manual for data entry. Some value fields restricted to control entry. Very little controls on entry into free form memo fields. The database incorporates data from other agencies for which we are not overly confident, but feel it does not affect the integrity of the data.
Future Development	**Proposed development path for the database:** **Extent the data are continuously being maintained:** 5-yearly reviews. Attempt to secure better data for Hutt Valley. **Certainty of the future of the data:** Data are well protected and intention is for databases to remain managed in perpetuity (or for as long as the RMAct and Biosecurity Act require)

Key Native Ecosystems (KNE) Geodatabase

Title	Key Native Ecosystems (KNE) Geodatabase
Abstract	This geodatabase is in the early stages of development. Like the wetlands database it consists of GIS data entry/reporting with Access database to hold non-GIS attribute data. The data include spatial extent of pest animal control, site assessment (dominant vegetation, species of indigenous plant and animal observed during assessment), GPS locations of observations (if available), and pest control history. In the future it is hoped this database will also hold monitoring information and casual rare/threatened species observations for the Wellington Region (subject to the lack of access to similar, national databases and $$). In the interim, the field assessment data are safely held in an Access database, and the spatial extent data are safely held as a GIS layer.
Purpose	Information warehouse for planning and reporting effective pest control
Contact Person	John Gibson (GIS), Fiona Bancroft (Biosecurity) / James Lambie (Biosecurity)
Organisation	Greater Wellington
Temporal Extent	1995 to present
Spatial Extent	Wellington Region
Number of Records	About 200 KNE site assessments.
Data Standards	**Rank:** 3–4 in terms of data collection, 2 in terms of data storage **Standards Used:** Largely follows DOC PNA site assessment **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.
Data Quality	**Rank:** 3 **Notes:** Largely follows DOC PNA site assessment, with statements to assist scoring
Future Development	**Proposed development path for the database:** KNE database to go capture much of the biosecurity biodata including casual observations. **Extent the data are continuously being maintained:** For KNE, every time a new KNE is added, plus every time a KNE undergoes pest management (sometimes annually). **Certainty of the future of the data:** Data are well protected and intention is for databases to remain managed in perpetuity (or for as long as the RMAct and Biosecurity Act require)

Parks and Forests Bird Count data

Title	Parks and Forests Bird Count data
Abstract	5-minute bird count data, and point-distance or transect – distance data on

	native birds in specific locations. Data stored in Excel spreadsheet. Observation points have GPS / Map grid reference.
Purpose	Response monitoring of pest control and long-term ecological monitoring
Contact Person	Philippa Crisp
Organisation	Greater Wellington
Temporal Extent	1996 to Present. Bulk of data 2001–2004
Spatial Extent	KNEs: Pounui KNE (South Wairarapa), Porirua Scenic Reserve (Porirua City), Keith George Memorial Reserve (Lower Hutt), Tauherenikau Bush KNE (Featherston), Parks: Wanuiomata, Akatarawa, East Harbour, Orongorongo and other GW lands.
Number of Records	Upwards of 15 000 records. Distance observations of tomtit, or tui/bellbird/warbler in each of the reserves named plus records of any other native species seen including kaka, käkäriki, whitehead and falcon.
Data Standards	**Rank:** 4–5 in terms of data collection, 1 in terms of records storage **Standards Used:** Follows recognised 5-minute and distance count methods **Notes:** Although we follow recognised protocol, we still collect data using three different methods. Add to this that there are two sets of Excel datasets, exemplifies the lack of standards here.
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.
Data Quality	**Rank:** 2 **Notes:** Follows recognised 5-minute and distance count methods. All staff trained. No controls on data entry.
Future Development	**Proposed development path for the database:** **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Biosecurity Bird Count data

Title	Biosecurity Bird Count data
Abstract	5-minute bird count data, and point-distance or transect – distance data on native birds in specific locations. Data stored in Excel spreadsheet. Observation points have GPS / Map grid reference.
Purpose	Response monitoring of pest control
Contact Person	James Lambie
Organisation	Greater Wellington
Temporal Extent	2004–2005

Spatial Extent	KNEs: Pounui KNE (South Wairarapa), Porirua Scenic Reserve (Porirua City), Keith George Memorial Reserve (Lower Hutt), Tauherenikau Bush KNE (Featherston), Parks: Wanuiomata, Akatarawa, East Harbour, Orongorongo and other GW lands.
Number of Records	Upwards of 15 000 records. Distance observations of tomtit, or tui/bellbird/warbler in each of the reserves named plus records of any other native species seen including kaka, käkäriki, whitehead and falcon.
Data Standards	**Rank:** 4–5 in terms of data collection, 1 in terms of records storage **Standards Used:** Follows recognised 5-minute and distance count methods **Notes:** Although we follow recognised protocol, we still collect data using three different methods. Add to this that there are two sets of Excel datasets, exemplifies the lack of standards here.
Data Dictionaries	**Dictionaries used:** None **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.
Data Quality	**Rank:** 2 **Notes:** Follows recognised 5-minute and distance count methods. All staff trained. No controls on data entry.
Future Development	**Proposed development path for the database:** **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Pest Plant extent in regional parks (Parks & Forests)

Title	Pest Plant extent in regional parks (Parks & Forests)
Abstract	GIS spatial extent layer of environmental weeds (one layer per weed) in specific reserves.
Purpose	Planning and reporting pest plant control.
Contact Person	Philippa Crisp
Organisation	Greater Wellington
Temporal Extent	2001 and 2003
Spatial Extent	GW lands
Number of Records	?? (thousands of observations) of approximately 400 naturalised species.
Data Standards	**Rank:** 3 **Standards Used:** Follows national plant pest recording and reporting guidelines

	Notes: The database is our own design but follows that used by others.
Data Dictionaries	**Dictionaries used:** Some internal restrictions on what can be entered into the species field (drop down prompt) **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.
Data Quality	**Rank:** 2 **Notes:** Training, forms and procedures well documented. Data entry controlled to some extent. Recent exercise to look at data integrity showed large short falls in this system.
Future Development	**Proposed development path for the database:** Pest plants database to be upgraded to facilitate extraction of data through GIS GUI and address recently recognised data integrity shortfalls **Extent the data are continuously being maintained:** For pest plants, annually / weekly maintenance. **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Pest Plant location and extent Geodatabase (Biosecurity)

Title	Pest Plant location and extent Geodatabase (Biosecurity)
Abstract	Impromptu database with GIS GUI for spatial extent, point location, and density information related to pest plants specified in Greater Wellington's RPMS.
Purpose	Planning and reporting pest plant control.
Contact Person	Dave Bayly / Mark McAlpine
Organisation	Greater Wellington
Temporal Extent	1995 to present
Spatial Extent	Wellington Region
Number of Records	?? (thousands of observations) of approximately 400 naturalised species.
Data Standards	**Rank:** 3 **Standards Used:** Follows national plant pest recording and reporting guidelines **Notes:** The database is our own design but follows that used by others.
Data Dictionaries	**Dictionaries used:** Some internal restrictions on what can be entered into the species field (drop down prompt) **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy /	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.

Procedures	
Data Quality	**Rank:** 2 **Notes:** Training, forms and procedures well documented. Data entry controlled to some extent. Recent exercise to look at data integrity showed large short falls in this system.
Future Development	**Proposed development path for the database:** Pest plants database to be upgraded to facilitate extraction of data through GIS GUI and address recently recognised data integrity shortfalls **Extent the data are continuously being maintained:** For pest plants, annually / weekly maintenance. **Certainty of the future of the data:** Data are well protected and intention is for databases to remain managed in perpetuity (or for as long as the RMAct and Biosecurity Act require)

Possum Residual Trap-catch data

Title	Possum Residual Trap-catch data
Abstract	Residual trap-catch lines results of possum monitoring.
Purpose	The bulk of the data are collected as part of auditing the effectiveness of possum control under the national bovine TB vector control programme. Some of the data are collected for planning future possum control operations or determining the current status of possum populations (i.e. data used to determine whether possum control is needed, or is present intensity of control is sufficient). The regional data also contributes to understanding the pressure the possum population may be placing on our highly prized indigenous forest ecosystems
Contact Person	James Lambie (Biosecurity)
Organisation	Greater Wellington
Temporal Extent	1/3/1995 (a early few records) to current
Spatial Extent	Wellington Region
Number of Records	11597 individual line results (as of 4/12/2004). The bulk of records (approx 6000) are of data collected since 1/7/2002.
Data Standards	**Rank:** 4 **Standards Used:** Follows NPCA protocol **Notes:** Data checking and data storage procedures is ISO 9001: 2000 approved.
Data Dictionaries	**Dictionaries used:** None **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	ISO 9001:2000 approved Quality Manual which includes data management policies. Fully complied with (audited by external agent).
Data Quality	**Rank:** 5 **Notes:** Follows NPCA protocol Data checking and data storage procedures is ISO 9001: 2000 approved. Double entry for data. Documented quality

	control checks. Quality of output checked by at least three different people (data entry operator, manager, client)
Future Development	**Proposed development path for the database:** **Extent the data are continuously being maintained:** Daily maintenance **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Parks & Forests rat/rodent and mustelid data

Title	Parks & Forests rat/rodent and mustelid data
Abstract	Rat/Rodent records from tracking tunnel monitoring, following DOC guidlelines.
Purpose	Understanding pest ecology and inferential statistics on rodent damage to forest ecosystems. Measure of effectiveness of rodent control/by-kill of rodents during other pest control activities.
Contact Person	Philippa Crisp
Organisation	Greater Wellington
Temporal Extent	2003 to present
Spatial Extent	Parks: Wauiniomata – Orongorongo, Akatarawa–Kaitoke, East Harbour, and other GW lands. KNEs: Haywards Scenic Reserve (Eastern Hutt), Porirua Scenic Reserve, Wrights Hill KNE (Wellington), Pounui KNE, Tauherenikau Bush KNE, Keith George Scenic Reserve, Fensham Reserve KNE (Forest and Bird – near Carterton), Witako KNE (Siverstream), Johnosonville Park (Johnsonville),
Number of Records	In excess of 350 individual tracking tunnel results.
Data Standards	**Rank:** 3 **Standards Used:** Follows DOC protocol for collection, entry and storage. **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.
Data Quality	**Rank:** 2 **Notes:** Standard Operating Procedure for data collection. No verification of data. Outliers checked to see they are not mistakes.
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:**

	Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Biosecurity rat/rodent data

Title	Biosecurity rat/rodent data
Abstract	Rat/Rodent records from tracking tunnel monitoring, following DOC guidlelines.
Purpose	Understanding pest ecology and inferential statistics on rodent damage to forest ecosystems. Measure of effectiveness of rodent control/by-kill of rodents during other pest control activities.
Contact Person	James Lambie/Sara Moylan
Organisation	Greater Wellington
Temporal Extent	2003 to present
Spatial Extent	Parks: Wauiniomata – Orongorongo, Akatarawa–Kaitoke, East Harbour, and other GW lands. KNEs: Haywards Scenic Reserve (Eastern Hutt), Porirua Scenic Reserve, Wrights Hill KNE (Wellington), Pounui KNE, Tauherenikau Bush KNE, Keith George Scenic Reserve, Fensham Reserve KNE (Forest and Bird – near Carterton), Witako KNE (Siverstream), Johnosonville Park (Johnsonville),
Number of Records	In excess of 350 individual tracking tunnel results.
Data Standards	**Rank:** 3 **Standards Used:** Follows DOC protocol for collection, entry and storage. **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.
Data Quality	**Rank:** 2 **Notes:** Standard Operating Procedure for data collection. No verification of data. Outliers checked to see they are not mistakes.
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Parks & Forests goat culling records

Title	Parks & Forests goat culling records
Abstract	Records of Goat culling in Wellington regional parks held in Excel database.
Purpose	Planning and reporting goat control
Contact Person	Philippa Crisp
Organisation	Greater Wellington
Temporal Extent	19?? To present
Spatial Extent	Specific regional parks.
Number of Records	unknown
Data Standards	**Rank:** 2 **Standards Used:** Internally defined, but follows much the same process as others **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	Informal or not written down.
Data Quality	**Rank:** 2 **Notes:** Assumes honesty in culling records. Assumes consensus in spatial layer is a realistic picture.
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Goat spatial extent GIS layer (Biosecurity)

Title	Goat spatial extent GIS layer (Biosecurity)
Abstract	Spatial extent layer of goat density for Wellington Region using 'professional consensus' approach (involving DOC and GW goat cullers).
Purpose	Planning and reporting goat control
Contact Person	James Lambie
Organisation	Greater Wellington
Temporal Extent	2000–2001 (one-off map)
Spatial Extent	Wellington Region
Number of Records	One-off GIS layer.

Data Standards	**Rank:** 2 **Standards Used:** **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	Informal or not written down.
Data Quality	**Rank:** 2 **Notes:**
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Goat spatial map at risk of deletion as data are becoming obsolete / not being used.

Various pest animal records not identified above

Title	Various pest animal records not identified above 1. Rabbit and possum trend monitoring (Biosecurity) 2. Rook location / trend monitoring 3. Magpie trend monitoring 4. Possum night counts
Abstract	Records of trend monitoring for RPMS pests and other pests of interest. All data held in Excel spreadsheet.
Purpose	Planning and reporting the effectiveness of pest control and pest density monitoring.
Contact Person	James Lambie / Ray Clarey (Biosecurity)
Organisation	Greater Wellington
Temporal Extent	1995 to present (except Magpies)
Spatial Extent	Rabbits: Two specific trend monitoring routes (One on Kapiti Coast and one in Northern Wairarpa) Rooks: All current rookeries (Wellington Reigon) Mapies: (Planned, yet to collect data) Wairarapa. Possums: One specific trend monitoring route in Belmont Regional Park.
Number of Records	5 Excel workbooks.
Data Standards	**Rank:** 3 for data collection, 1–2 for data entry and storage **Standards Used:** Follows industry recognised protocols for data collection. Some controls, but not explicitly recorded. **Notes:** Formality of data entry not required. Low volume data allows for

	double checking (data entry and analysis mistakes are obvious).
Data Dictionaries	**Dictionaries used:** None **Notes:** Unnecessary for RPMS pests, as pest name is easily made to be congruent with regional nomenclature (Ques 6)
Data Management Policy / Procedures	Regional Pest Management Strategy (RPMS) documents which names will be used in pest plant and animals databases.
Data Quality	**Rank:** 2 **Notes:** Follows industry recognised protocols for data collection. Some controls, but not explicitly recorded. Formality of data entry not required. Low volume data allows for double checking (data entry and analysis mistakes are obvious).
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Various special (e.g. rata and rare/endangered) plant species records / Casual observations of rare/endangered animalia (including invertebrates)

Title	Various special (e.g. rata and rare/endangered) plant species records / Casual observations of rare/endangered animalia (including invertebrates)
Abstract	Collections of formal and informal observations of species of interest.
Purpose	Formal records: Reporting of ecological health of forests Casual: Interest.
Contact Person	Philippa Crisp (Parks & Forests), (many) (Biosecurity).
Organisation	Greater Wellington
Temporal Extent	1970 (Parks) to present (all)
Spatial Extent	Wellington Region
Number of Records	Parks: around 4000 records Other observers: ??
Data Standards	**Rank:** 1 **Standards Used:** **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	Informal or not written down.
Data Quality	**Rank:** 1

	Notes: Depends largely on the habits and training of individual observers.
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Fish and Macro-invertebrates

Title	Fish and Macro-invertebrates
Abstract	All records of Marco invertebrate monitoring is stored on formal database. All fish records go straight into FBIS.
Purpose	Water quality monitoring and ecosystem health indicators.
Contact Person	Alton Perrie
Organisation	Greater Wellington
Temporal Extent	1980s?? to present. Bulk of data since early 2000s
Spatial Extent	At present; 62 sites across the Wellington Region
Number of Records	??
Data Standards	**Rank:** 4 **Standards Used:** Follows national protocols. **Notes:** FBIS data toward the higher end. Still potential risk for wayward data entry.
Data Dictionaries	**Dictionaries used:** Not known by me, but presumed to use national databases **Notes:**
Data Management Policy / Procedures	Not known by me, but assumed to have some policies. This Dept explored the idea of an ISO-accredited system but has yet to implement it.
Data Quality	**Rank:** 3 **Notes:** I assume the data collection and entry follows national protocols.
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

Phenology data

Title	Phenology data
Abstract	Records of flowering/Fruiting cycles
Purpose	To keep a check on the health of rare species. Ecological investigation/research.
Contact Person	Philippa Crisp
Organisation	Greater Wellington
Temporal Extent	Monthly from 2001 to 2004
Spatial Extent	GW lands
Number of Records	Around 3000 records
Data Standards	**Rank:** 2 **Standards Used:** Internal guidelines **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	Informal or not written down.
Data Quality	**Rank:** 3 **Notes:**
Future Development	**Proposed development path for the database:** Uncertain **Extent the data are continuously being maintained:** Uncertain **Certainty of the future of the data:** Reasonably safe. The maintenance of much of the data is implied in the job description of the custodian.

KNE Invertebrate Monitoring

Title	KNE Invertebrate Monitoring
Abstract	Invertebrate pitfall trapping (using DOC guidelines) records for specific KNEs, held in Excel spreadsheet. Invert identified by size class and to family level (and species where feasible)
Purpose	Monitoring ecological health of KNEs.
Contact Person	James Lambie
Organisation	Greater Wellington
Temporal Extent	2002 to present
Spatial Extent	East Harbour, Haywards Scenic Reserve, Porirua Scenic Reserve, Wrights Hill KNE, Pounui KNE, Keith George Scenic Reserve, Fensham Reserve KNE, Witako KNE, Johnosonville Park.
Number of Records	Approximately 540 individual pitfall trap records. Approximately 70 morpho-species in reference collection.

Data Standards	**Rank:** 1 **Standards Used:** Data collection follows DOC protocol. **Notes:** Internal Storage our own design and not particularly strong controls on data entry
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	Informal or not written down.
Data Quality	**Rank:** 3 collection, 1 storage **Notes:** Data collection follows DOC protocol. Sample analysis done by two technically competent staff, Competence left largely to trust. Not particularly strong controls on data entry.
Future Development	**Proposed development path for the database:** KNE database to go capture much of the biosecurity biodata including casual observations. **Extent the data are continuously being maintained:** Additions to data as changes in the spatial extent of object of interest occurs. **Certainty of the future of the data:** Data are well protected and intention is for databases to remain managed in perpetuity (or for as long as the RMAct and Biosecurity Act require)

Development, Integration and Interoperability

The following are responses by Greater Wellington to questions on future biodiversity database developments and integration/interoperability issues.

Approach to Metadata Management	All spatial extent data listed on the main GW GIS sever requires geo-referenced location, technology use to collect data, scale, a little on who collected it and why. These data are listed on the staff intranet to aid GIS layer searches. Our GIS is at a point that it requires quite rigorous needs for metadata, and is usually well managed and adhered to. Data includes: Title, Abstract, Author, date of publication, reliability, attribute information, copyright and where to find the spatial layer.
New Databases Planned	There is the potential for the need for a coastal habitats database. Creation of such a database will only come about if we cannot easily access / cannot rely other coastal datasets. KNE Geodatabase – plans to expand to incorporate much of the biosecurity biodata are still in their infancy.
Hardware and Software Platforms	Excel, Access, Arc Map v9.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Integration will probably be only by way of 'cutting and pasting' into other data sets e.g. pest Plants database information was copied straight into Bioweb. We are aware of the challenges of concurrency and the risk of double counting the same entity, but have no formal policies in place to

	avoid this. This would be a reasonable benchmark for other datasets – i.e. we could do wholesale handover of data, but anything else would be tricky. Some data we are collecting goes directly into national databases. **Other fields that might be of benefit to major biodatabases:** Wetlands: Vegetation (structure, species from most to least dominant – top 5), assessment of its quality (incl. count of no of culverts etc.). Largely, our additional fields such as pest control history for certain areas of special habitat are outside the scope of TFBIS?
Barriers to Integration with Major Biodatabases	Cultural (people won't want to share data) Trust (won't want to share with some people because they won't use it appropriately), some legal (private land data), technical connectivity (in terms of the tools we have and the lack of people who know how to use them).
Benefits of Integration with Major Biodatabases	To Greater Wellington – the fact that other people make observations in our region, so we'll get to see more data. To know what our neighbors are up to in terms of biosecurity, for border control, or differences in policy. Assurances of custodianship of the data (e.g. if NIWA is looking after it). Good habits, adoption of better practice than we would have achieved ourselves. To others – similar, more data (e.g. save OSNZ time in collecting their own records), public could have the opportunity to save their data somewhere.
Externally Stored Data	**Data currently stored in external organisation's systems:** FBIS, NVS, DOC Threatened species database (mistletoe records), NZAC, Fungal Herbaria, Allen Herbarium (moss flora)
Integration with other Regional Councils	**Benefits of data integration:** As for question 14, not much happening yet. **Benefits of sharing systems/infrastructure:** Some districts within the region that presently don't have the technological/knowledge capacity would be able to better participate in regional decision making. With other regions, it could avoid duplication. **Examples of unintentional or unavoidable reinvention of the wheel:** ecoBase and KNE are fairly similar, both were set up at roughly the same time, Horizons had thought about how to structure before collecting it, GW had collected it first. Birds data could be a duplication of effort as other people are sampling the same areas, and also duplication of databases and database design effort as there is no national birds database.

(7) Horizons Regional Council

Natural Areas Database

Title	Natural Areas Database
Abstract	The Natural Areas Database is a part of the Incidents Database. It contains management information, understorey plot data and Foliar Browse Index data from field surveys of natural areas outside DOC land.
Purpose	Assist with pest management and biodiversity management functions under the Biosecurity and Resource Mananagement Acts
Contact Person	Malcolm Todd
Organisation	Horizons Regional Council
Temporal Extent	1998 to present

Spatial Extent	Manawatu Wanganui region plus a buffer zone. Data points from PNA survey for Wanganui Conservancy etc.
Number of Records	296 natural areas entered, 29 have FBI plot, viewpoint FBI plot or understorey plot data.
Approach to Metadata Management	Databases were designed by Datacom in Hamilton. We have user specs, design specs etc. etc.
Data Standards	**Rank:** 3 **Standards Used:** **Notes:** Don't know, probably 5
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	
Data Quality	**Rank:** 3 **Notes:**
Future Development	**Proposed development path for the database:** Don't have one **Extent the data are continuously being maintained:** Small amount **Certainty of the future of the data:** Not certain.

ecoBase

Title	ecoBase
Abstract	ecoBase records biodiversity data for whole region collected by Horizons, DOC and others. It is used to determine what areas require active management as a priority. The Natural Areas database then records monitoring information on those particular sites. ecoBase contains historical biodiversity survey data, PNA survey data and information from DOC. Reference points and dates. Rapid ecological assessment to give regional overview, inventory of biodiversity sites for prioritising for management. Assessments done for wetlands but not forest reference. Linking up with distance sampling software developed overseas for bird monitoring (more rigour, than 5-minute bird counts, distance from a line).
Purpose	Priorise research, monitoring, guidance for prioritisation of areas to restore, and to provide policy advice for district councils.
Contact Person	Helmut Janssen
Organisation	Horizons Regional Council
Temporal Extent	1908 to 2005. 80s and 90s PNA surveys and more data.
Spatial Extent	Manawatu Wanganui region plus a buffer zone. Data points from PNA survey for Wanganui Conservancy etc.
Number of Records	2533 site records. Links with Names.

Approach to Metadata Management	There are written reports on ecoBase that describe what's in it.
Data Standards	**Rank:** 5 **Standards Used:** Darwin Core **Notes:** Started trying to develop our standard with DOC but didn't get involvement from them so developed our own. Went to the TFBIS workshop and had our eyes opened to international data exchange standards. Data are now exportable in Darwin Core. Very flexible and well set up
Data Dictionaries	**Dictionaries used:** Landcare Research Plant Names Database **Notes:** Link internally to NVS datapoints but not data dictionary style.
Data Management Policy / Procedures	Looked into it. ecoBase is a standalone Access DB, project lined up sometime this year to put it into SQL Server. Will bring it in line with the council's data management processes.
Data Quality	**Rank:** 3 **Notes:** Using REA descriptors like PNA. There is a lot more detail we're putting in. Subdivided sites into many habitats, and put threats to the site. People collect data, variability in quality for new people, no certainty statements defined. So no double entry or proofing. Not used for scientific purposes so rigorous checks not necessary. Plot entry data with data entry wizards will rank a 3.
Future Development	**Proposed development path for the database:** It is a portable database currently. We plan to integrate it into the main council databases and make some of the data web accessible. **Extent the data are continuously being maintained:** Continuously used, REA data being entered. **Certainty of the future of the data:** Reasonably certain but don't know what would happen if I left the organisation.

Development, Integration and Interoperability

The following are responses by Horizons Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	None
Hardware and Software Platforms	Natural Areas Database is inside the Incidents database which is an SQL database with Powerbuilder front end running on Windows 2000. ecoBase is an Access database.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Natural Areas Database: Good. We don't hold specimens but have all the other data ecoBase: Would be very good for taxon name, location and date. We are using and aware of standard exchange schemas

	Other fields that might be of benefit to major biodatabases: Natural Areas Database: Management Info: info on the whole site: Owner, owner attitude to retirement and pest control owner actions, regional council actions, fence status, grazing status, pests and threat, legal protection and type. All the fields that get recorded for FBI plots; plot info: slope, gridref, landform, vegetation, altitude, pest species (weeds and animals). And the tree info: Species, dbh, distance and direction from plot centre, stem/tree, commonness, canopy/understorey/emergent/isolated tree, dieback top/whole, browse top/whole, canopy density, stem use, flowering & fruiting. FBI viewpoint plots get a subset of the FBI plot info, but for more individuals. Understorey plots: recording species and height within 49-cm radius of peg, up to 1.35 m, within five height classes. ecoBase: Habitat types (50), structure (how high, levels of species), estimates of abundance (counts or groups or % depending on size of habitats), threats (animal, human) – all part of rapid ecological assessment. For plot data, all from forest monitoring toolkit for transects etc.
Barriers to Integration with Major Biodatabases	Funding and time
Benefits of Integration with Major Biodatabases	Malcolm Todd: To our organisation – would give us context for how well were doing, and open possibility of getting results of research on our data. For others – same, plus larger datasets Helmut Janssen: In the future I see our database being a node within a much larger system. We collect data at our scale, feed it up and access data at neighboring regions. Biodiversity doesn't stop at our boundaries, we're a part of a much larger system. In terms of benefit to major database holders – much more data collection, access to other data for research purposes. It also has implications for data quality. Integration encourages standards adoption which in turn encourages better data collection practices in regional councils and other organisations such as NZERN.
Externally Stored Data	**Data currently stored in external organisation's systems:** Malcolm Todd: Not me, others use NIWA Freshwater fish DBase Helmut Janssen: No, but that is the ultimate aim
Integration with other Regional Councils	**Benefits of data integration:** There would be benefits in terms of the fact that we all deal with similar problems. Exchanging data would also lead to the exchange of knowledge about good data management practices, and also knowledge about resource management. Data are potentially a carrier wave for other knowledge. There is not yet enough cross communication between councils in terms of what can be done, and how to develop things further. It is improving but it is based on the people involved. There is no means of doing this consistently or nationally. At the moment leadership is very dependent on the person driving biodiversity management in the organisation. Good biodiversity data management isn't institutionalised yet. There is a high risk of loss of knowledge and capacity if people leave. **Benefits of sharing systems/infrastructure:** Peer-o-peer-style integration of data perhaps. Fairly independent systems linking up and exchanging data rather than a large centralised system.

	Examples of unintentional or unavoidable reinvention of the wheel: Malcolm Todd: Of course! NADbase overlaps a little with ecoBase within our own organisation, then we have overlap with Peter Handfords FORMAK Dbase and I'm sure there are others. Helmut Janssen: Think there might be overlaps with Bioweb, had a lot of difficulty getting a response from DOC when we tried to communicate with them when we were developing ecoBase.

(8) Marlborough District Council

SIGAreas

Title	SIGAreas
Abstract	Database of significant natural areas on private property – what type of communities, significant individual species, how representative or rare.
Purpose	To classify/define what levels of protection are required.
Contact Person	Peter Hamill
Organisation	Marlborough District Council
Temporal Extent	Started in 2002
Spatial Extent	Just Marlborough
Number of Records	400 sites
Data Standards	**Rank:** 2 **Standards Used:** Our own standards **Notes:** Look at rareness, intactness, list of criteria we've developed
Data Dictionaries	**Dictionaries used:** **Notes:**
Data Management Policy / Procedures	Data entry done only by one or two people so no documented procedures
Data Quality	**Rank:** **Notes:**
Future Development	**Proposed development path for the database:** Not broke so don't fix it, doing fine. **Extent the data are continuously being maintained:** Continuously used, REA data being entered. **Certainty of the future of the data:** Well cared for, very well used.

QDAS

Title	QDAS
Abstract	Freshwater quality data, macroinvetebrates, all species and relative abundance
Purpose	
Contact Person	Peter Hamill

Organisation	Marlborough District Council
Temporal Extent	Since year 2000
Spatial Extent	Marlborough Water ways
Number of Records	20 sites twice a year
Data Standards	**Rank:** 4 **Standards Used:** Standard MCI **Notes:**
Data Dictionaries	**Dictionaries used:** Index given to us from Cawthron Institute **Notes:** This was a once off batch process. Now if a new species pops up we add it in manually.
Data Management Policy / Procedures	Documented manuals for data entry processes
Data Quality	**Rank:** 4 **Notes:** Once data are entered it is printed out, checked by another person, and assigned a confidence level. Information back from laboratories is ISO accredited. No specific data integrity measures.
Future Development	**Proposed development path for the database:** Not broke so don't fix it, doing fine. **Extent the data are continuously being maintained:** **Certainty of the future of the data:** Well cared for, very well used.

The following are responses by Marlborough Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	
Hardware and Software Platforms	Microsoft Access. Some council databases are being put on to SQL Server but not yet for this one.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** **Other fields that might be of benefit to major biodatabases:**
Barriers to Integration with Major Biodatabases	
Benefits of Integration with Major Biodatabases	
Externally Stored Data	**Data currently stored in external organisation's systems:**
Integration with other Regional	**Benefits of data integration:** **Benefits of sharing systems/infrastructure:**

Councils	Examples of unintentional or unavoidable reinvention of the wheel:

Ecodata

Title	Ecodata
Abstract	Database started in the marine area. There is lots of scientific research done in the Marlborough Sounds, much by science providers to MRC. Ecodata records some details on the research that's been done, in particular the rare or unusual species in that location.
Purpose	
Contact Person	Peter Hamill
Organisation	Marlborough District Council
Temporal Extent	Research done since 1994
Spatial Extent	Only Marlborough Sounds area
Number of Records	3000 site records – 50 would be freshwater or terrestrial
Data Standards	**Rank:** 2 **Standards Used:** Our own standards **Notes:** Experts make a call on what are significant species
Data Dictionaries	**Dictionaries used:** Internal – species glossary lookup **Notes:** Provided once off and has grown from there
Data Management Policy / Procedures	Data entry done only by one or two people so no documented procedures
Data Quality	**Rank:** 2 **Notes:** Just summaries of reports etc. so no quality system
Future Development	**Proposed development path for the database:** Ecodata are in development mode currently. Already mapped within science providers in Marlborough region, we're running it internally here. Want to make it accessible over the Web. **Extent the data are continuously being maintained:** Ongoing based on new research. Making it Web accessible will mean research contractors can enter data themselves. **Certainty of the future of the data:** Well cared for, very well used.

The following are responses by Marlborough Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	
Hardware and Software Platforms	Microsoft Access. Some council databases are being put on to SQL Server but not yet for this one.
Integration with Major	**Ability to provide or exchange very basic collection or observation data:**

Biodatabases	**Other fields that might be of benefit to major biodatabases:**
Barriers to Integration with Major Biodatabases	
Benefits of Integration with Major Biodatabases	
Externally Stored Data	**Data currently stored in external organisation's systems:**
Integration with other Regional Councils	**Benefits of data integration:** **Benefits of sharing systems/infrastructure:** **Examples of unintentional or unavoidable reinvention of the wheel:**

Development, Integration and Interoperability

The following are responses by Marlborough District Council to questions on future biodiversity database developments and integration/interoperability issues.

Approach to Metadata Management	None, people generally know what datasets we have.
New Databases Planned	None apart from a small photographic database of macroinvetebrate pictures
Hardware and Software Platforms	Microsoft Access. Some council databases are being put on to SQL Server but not yet for this one.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Our fish data goes in to FBIS. I worry about misidentification, the level of expertise varies so much and I wonder that a lot of stuff that goes into FBIS is wrong. I could get the genus but would never attempt to get the species, just wouldn't have the expertise. Macroinvetebrates data – no worries about that, if there was a national database they'd be entering it themselves (the researchers we contract). **Other fields that might be of benefit to major biodatabases:** Abundance, significant areas – level of rareness, how robust the site is (ranking), Geoff Walls – one of the better ecologists does this for this. Representativeness, rarity, diversity and pattern, distinctiveness, size & shape, connectivity, sustainability. Each ranked individually, then an overall significant rank of the site.
Barriers to Integration with Major Biodatabases	Time vs value – would be low priority in terms of national vs our region. Don't think it should be the ratepayers of our region paying for it. Funding might not help because we might not be able to get staffing. Unless its something like fish database where there are other people entering in the data, so we don't have to store our own data at all. Won't be anyone else hosting freshwater macroinvetebrates data in Marlborough though.
Benefits of Integration with	If there was value, if everything was set up and ready to go, but don't want to double entry into our stuff at all. Unless it's been going for a while we

Major Biodatabases	don't want to be on the cutting edge of this stuff. We're a small council with a tiny amount of rate payers. Would follow after the big regional councils.
Externally Stored Data	**Data currently stored in external organisation's systems:** NZFFD is the only one.
Integration with other Regional Councils	**Benefits of data integration:** Can't see any immediate benefit apart from reinventing the wheel if we want to do something new. **Benefits of sharing systems/infrastructure:** Huge nightmare, wouldn't want to have anything to do with it. **Examples of unintentional or unavoidable reinvention of the wheel:** None.

(9) Nelson City Council

Nelson City Council did not provide a full response; the following are notes from a short phone conversation with Paul Sheldon:

For areas identified as significant conservation value we contracted Mike Harding to survey and prepare species lists, condition statements, define extent of these. Have a record for each of these sites, with the intention to review every five years. Mike is extending the survey by looking at new areas. This year he'll start surveying new sites. Data are currently sitting in a series of field sheets and electronic records of these. Had a quick look at ecoBase, thought that's a nice idea to do one day. Currently trying to get water quality, contaminated sites etc. done. Jointly looking at bringing in MonitoringPro with Trevor Jones at TDC. Our hydrology archive sits within TDC, pest management functions are managed by TDC, we try to make sure the monitoring we do, the tests, the way we report is compatible. Have the same engine for databases. Find all sorts of issues in the changes of the tests done over the years, quality of information.

(10) Northland Regional Council

Sites

Title	Sites
Abstract	Site specific SOE (state of the environment) monitoring, water quality monitoring, some macro-invertebrate data
Purpose	
Contact Person	Lisa Maria
Organisation	Northland Regional Council
Temporal Extent	
Spatial Extent	Northland
Number of Records	
Data Standards	**Rank:** 1 **Standards Used:** **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management	ISO & New Zealand GeoSpatial Metadata Standard – minimal

Policy / Procedures	
Data Quality	**Rank:** 3 **Notes:**
Future Development	**Proposed development path for the database:** Minimal **Extent the data are continuously being maintained:** Constant **Certainty of the future of the data:** Incorporated into main Regional Council data management system so no risk of loss

Biosecurity

Title	Biosecurity
Abstract	Biosecurity–pests (insects, animals and plants). Site information and pest density/population data
Purpose	To monitor the operations of the biosecurity department and the incursion, spread and control of pests (animals, plants, insects)
Contact Person	Lisa Maria
Organisation	Northland Regional Council
Temporal Extent	1984–
Spatial Extent	Northland
Number of Records	
Data Standards	**Rank:** 4 **Standards Used:** FORMAK (MfE) **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	ISO & New Zealand GeoSpatial Metadata Standard – minimal
Data Quality	**Rank:** 3 **Notes:**
Future Development	**Proposed development path for the database:** Minimal **Extent the data are continuously being maintained:** Constant **Certainty of the future of the data:** Incorporated into main Regional Council data management system so no risk of loss

Pest Eradication Contracts

Title	Pest Eradication Contracts

Abstract	Animal pest eradication contract coverages. Quality control and sampling locations for pest eradication contract areas.
Purpose	To monitor the operations of the contractors engaged in the management of pest eradication as well as records areas treated or under contract.
Contact Person	Lisa Maria
Organisation	Northland Regional Council
Temporal Extent	
Spatial Extent	Northland
Number of Records	
Data Standards	**Rank:** 1 **Standards Used:** **Notes:**
Data Dictionaries	**Dictionaries used:** None **Notes:**
Data Management Policy / Procedures	ISO & New Zealand GeoSpatial Metadata Standard – minimal
Data Quality	**Rank:** 3 **Notes:**
Future Development	**Proposed development path for the database:** Minimal **Extent the data are continuously being maintained:** Constant **Certainty of the future of the data:** This database may disappear because changes in the regional strategy mean we won't be recording the same things.

The following are responses by Northland Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	
Hardware and Software Platforms	
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** **Other fields that might be of benefit to major biodatabases:**
Barriers to Integration with Major Biodatabases	
Benefits of Integration with Major Biodatabases	

Externally Stored Data	**Data currently stored in external organisation's systems:**
Integration with other Regional Councils	**Benefits of data integration:** **Benefits of sharing systems/infrastructure:** **Examples of unintentional or unavoidable reinvention of the wheel:**

Development, Integration and Interoperability

The following are responses by Northland Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

Approach to Metadata Management	None
New Databases Planned	Wetland monitoring based on WONI. Lakes monitoring (biodiversity, biosecurity and water quality information) using NIWA standards (LakeSPI) and ANZECC water quality guidelines. Estuarine monitoring using Estuarine Monitoring Protocol (Cawthron Inst.).
Hardware and Software Platforms	Pathway/WorkSmart (GEAC), ArcGIS (ESRI), Microsoft SQL Server and O/S
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Able to provide that data, although the definition of integration would need to be clarified before meaningful responses could be given. **Other fields that might be of benefit to major biodatabases:** For the current dataset only basic density/population data but will include ecological monitoring data based on a national system developed by MfE (FORMAK). For future biodatabases we will collect data that follow national protocols and therefore will be of benefit to national databases
Barriers to Integration with Major Biodatabases	Data exchange formats/standards. Data consistency, conformity, accuracy, reliability, timeliness, etc. Security/Confidentiality of information.
Benefits of Integration with Major Biodatabases	Assessment of trends or biodiversity health nationally. May help funding bids locally if could identify nationally significant things in Northland.
Externally Stored Data	**Data currently stored in external organisation's systems:** No
Integration with other Regional Councils	**Benefits of data integration:** Biosecurity information from neighbouring Councils to undertake risk assessments to our region. Very limited use of other Council biodiversity data but that information is likely to be accessed in other ways – not via database exchange. **Benefits of sharing systems/infrastructure:** Limited – no real benefit perceived. **Examples of unintentional or unavoidable reinvention of the wheel:** No obvious examples

(11) Otago Regional Council

Biomonitoring database

Title	Biomonitoring database
Abstract	
Purpose	Maintain database of state of the environment stream macroinvetebrate and algal information.
Contact Person	Chris Arbuckle & Dave Turner, Environmental Monitoring Officer
Organisation	Otago Regional Council
Temporal Extent	1996 – present (2000 onwards data collected under new protocols)
Spatial Extent	Otago
Number of Records	Stream/site specific, 80 sites, measured annually
Approach to Metadata Management	All data collection (hydrometric and water quality) is ISO accredited, biodata aren't. Every site has full metadata per site, of the kind of data for that site. Water quality database has references to these site cards. We have a biodiversity officer, there are a number of sites under covenant. We have an active water plan, all wetlands and sites with certain values are mapped. We do a lot of instream habitat mapping, kept in reports, core data are in NZFFD, supplementary is not. We are trying to link the consents and state of environment monitoring databases together.
Data Standards	**Rank:** 4 **Standards Used:** Our own field protocol manuals. National protocol for sampling macroinvertebrates in wadeable streams. **Notes:** Defines protocol for collection of samples. MCI, species richness, others (don't rely soley on MCI). Same for algal sampling.
Data Dictionaries	**Dictionaries used:** Documents on fisheries, changing classification on our native fish in Otago, DOC native recovery programme (non migratory species). **Notes:** Maintaining own data dictionaries, double check against NZFFD. Very sure of what we're looking at as we're local people.
Data Management Policy / Procedures	Have a full in house data management manual. Have a dedicated internal data manager.
Data Quality	**Rank:** 4 **Notes:** Contract out the identification of invertebrates. Full quality control procedure, 10% of samples analysed are sent off again and reanalysed as a quality check. Could make a blanket certainty statement for the data (from 2000 onwards). Data pre 2000 is variable in quality in terms of how the metrics were calculated (would have to go back to the raw data and rework it). Will go for ISO accreditation for our data soon.
Future Development	**Proposed development path for the database:** Global dataset for the whole of council. Delivering hydrometric data on the Web now, intend to make all our datasets available. In full review at the moment, June 2006 to begin to transfer data. **Extent the data are continuously being maintained:**

	Annually, may move to biannually **Certainty of the future of the data:** Highly certain, collected for a purpose now

Development, Integration and Interoperability

The following are responses by Otago Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	May do mapping projects for each of our estuaries, and probably do a better job of wetlands in the medium term. Compliance officers currently maintain database of pest species (plant and animal), e.g. maps of gorse spreading, will become a full database in the fullness of time.
Hardware and Software Platforms	Citrix based, currently stored in Access Databases, currently evaluating platforms for full dataset. Mapinfo GIS tools, not used very powerfully here but that is changing.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Relatively easy standards wise and technically for 2000 onwards, in batch process. In the future may be able to connect in real time. **Other fields that might be of benefit to major biodatabases:** MCI, total richness, SQMCI (quantitive assessment), EPT, some with total numbers, water quality data (bi-monthly, so temporarily more detailed). Chris did a full stream database at Otago University. Most university data are never really published as it is done for specific outputs (student thesis).
Barriers to Integration with Major Biodatabases	Only technical connectivity and that is improving. Maybe some ownership issue but not a major problem with state of the environment data. Funding could be a barrier too if it is outside the remit of the rates collected for regional purposes.
Benefits of Integration with Major Biodatabases	No major perceived regional value. No other district councils do sampling of this nature. Some value in having it looked at in other ways, by other people, might lead to filling in some holes we don't we have. Value externally in terms of regional comparison, of scientific interest.
Externally Stored Data	**Data currently stored in external organisation's systems:** NZFFD, NIWA algal monitoring database. WONI (Wetlands of National Importance), held by DOC.
Integration with other Regional Councils	**Benefits of data integration:** Yes, it would be useful for comparison reasons. Share data now with Environment Southland and ECan for hydrometric data. **Benefits of sharing systems/infrastructure:** Of value in reducing cost, looked at this with Scott Crawford at Environment Southland. There are some challenges politically. **Examples of unintentional or unavoidable reinvention of the wheel:** More with major providers than other regional councils. We proposed building a complete Otago region fisheries database because we knew there was data in Otago Uni and DOC that wasn't in national databases. This resulted in those data getting in to the national databases and us not having to build the database. Landcare Research land cover database and LENZ seem to duplicate components. Needs to be more direction to councils as to

	what they might use those databases for. OBI process may help but there are constraints in terms of councils being reluctant to support things that will favour one rate payer over the other. Often the information on what you shouldn't use data for isn't provided in the end products and that can lead to errors. It's easy getting the data in there, much harder getting the metadata/information/knowledge in there. As soon as you get spatial data in you get far more risk of misinterpretation being made.

(12) Taranaki Regional Council

Freshwater Biological Database

Title	Freshwater Biological Database
Abstract	Samples of invertebrates, plants, algae and fish taken from rivers and lakes in the Taranaki region. Water temperatures, riparian cover (presence absence), river hydrology.
Purpose	Used for consent monitoring purposes, general investigations, and state of the environment monitoring.
Contact Person	Chris Fowles
Organisation	Taranaki Regional Council
Temporal Extent	Samples going back to 1980s, sites change over time.
Spatial Extent	Whole of Taranaki region
Number of Records	100s to 1000s of sites, just under 8000 samples
Approach to Metadata Management	Just in our heads.
Data Standards	**Rank:** 3 **Standards Used:** NZ Datum, 1947 standard. John Stark biological index for stony streams **Notes:**
Data Dictionaries	**Dictionaries used:** Tables of species names, Topographical maps for naming river systems **Notes:** Species names from published literature. Keep ourselves up to date by reviewing the literature periodically. Using topographic maps for river names. Also using our own interpretation/naming of individual sites.
Data Management Policy / Procedures	No
Data Quality	**Rank:** 4 **Notes:** Do have quality control systems in place. Adopted NZ freshwater protocols. We share samples with other acknowledged experts and have our IDs checked and written up as internal reports. No certainty statements but in any reporting we make internal or external, there is a comment about our protocols, sampling methods.
Future Development	**Proposed development path for the database:** Is currently on our openVMS, in Powerhouse, moving to SQL-Server in Powerbuilder on Windows

	Extent the data are continuously being maintained: Continuously entered on an almost daily basis. Data are used all the time for consent monitoring reports, used annually for internal state of the environment reporting, and in five yearly published state of the environment reports. **Certainty of the future of the data:** Very certain

Development, Integration and Interoperability

The following are responses by Taranaki Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	Will be doing one with Pest Plants to record them geospatially – just started planning this
Hardware and Software Platforms	Microsoft SQL-Server on Windows, Powerbuilder. System is very flexible, can be adapted to do many things. Biological monitoring was seen as a valuable tool very early on in Taranaki (20 years ago). Probably one of the biggest databases in NZ outside of research organisations. When the database was first set up 10 years ago we made the decision to digitise all historical data (where it could be accurately verified and was up to an acceptable standard). As an example of the benefit of this, we had a local iwi here who got interested in farming freshwater crayfish. They were considering the feasibility of this at a council meeting and turned around and asked if anyone knew where you actually found freshwater crayfish in the region. We were able in about ten minutes flat to do a query and get the exact site locations for all occurrences of freshwater crayfish in Taranaki. We are part of a group of regional councils, development work done by Datacom in Hamilton. Consents, Contacts, Pollution instance. If anyone changes anything others can have it, or anyone causes a new module to be developed they can have it at 10% of development cost. Horizons, Otago, Southland, West Coast, Environment Waikato, Taranaki. Mainly us smaller ones banded together for financial reasons. The Freshwater Biological database (includes groundwater, marine etc. too) fits under this technology paradigm, and may be taken up by the other RCs in this group.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Could copy over but not connect in real time. Have sent quite a lot of data to various consultancies (including Cawthron) in the past for differing purposes. We retain all our specimens here. No IP issues. **Other fields that might be of benefit to major biodatabases:** MCI, how many found (relative abundance), sampling method, environment it was found in (water temperature, substrate type, abundance of periphyton, riparian vegetation cover), site (lat long location, height above sea level, distance to mouth of river), hydrological information (river flows etc.). Chemical data can be provided through link by site name to our chemical database.
Barriers to Integration with Major Biodatabases	Batches provided on request, different data standards (e.g. most rivers in Taranaki are off the mountain so fairly unique). We score differently to the national systems as our scoring methods have been modified/refined from original MCI methods. These are completely documented in our

	publications. There are a few additional taxa we have found that others may have not.
Benefits of Integration with Major Biodatabases	National basis, MfE purposes for SoE monitoring. Important to ensure standards are aligned however. Important also to ensure local knowledge is retained/transferred when sharing our data outside the region. Also benefits in comparing us with other areas of the country.
Externally Stored Data	**Data currently stored in external organisation's systems:** NZFFD
Integration with other Regional Councils	**Benefits of data integration:** Comparing us with other areas of the country **Benefits of sharing systems/infrastructure:** Decentralised copies of systems are valuable to us as a smaller regional council, would look seriously at new centralised systems (to save us having to develop and maintain ourselves) **Examples of unintentional or unavoidable reinvention of the wheel:** No. Don't know what other councils have got. Know that this system we have other councils don't have. Other councils put different emphasis on biological monitoring (normally due to enthusiasm of individuals). Also differences based on geography. Have had a lot of queries about our macroinvertebrate database early on. Now it seems that others are developing their own. We think ours is fairly unique and of high quality both in terms of the data, and of the database system itself.

(13) Tasman District Council

Tasman District Council did not provide a full response; however, the following are notes from a short phone conversation with Peter Inwood:

Have a wetlands database, being revised at the moment, water bodies database initiated by fish & game – ecological information, building an environmental monitoring dataset – water sampling, pulling in 6 or 7 Access Databases and information out of box files etc., and batch process to load water testing data etc. from Cawthron. Rob Smith, Trevor James doing some work on the invertebrate area. Have over 300 datasets fully documented in our GIS system etc. We have a regional GIS initiative in top of the South, meet every 6 weeks or so. Involves electricity company, city council etc.

(14) West Coast Regional Council

Macroinveterbrates Spreadsheet

Title	Macroinveterbrates Spreadsheet
Abstract	Samples collected for the purpose of state of the environment monitoring, and for specific studies. Long term datasets, collected once a year. State of the environment monitoring sites change over time for various reasons. We collect the samples, Cawthron sorts and identifies down to MCI (Macroinvetebrate Community Index) level. No abundance data, does have diversity indicies.
Purpose	To monitor water quality
Contact Person	Jonny Horrocks
Organisation	West Coast Regional Council
Temporal Extent	Some from the mid 90s, some more recent. Sampled annually in Autumn.

Spatial Extent	West Coast region (mainly areas that aren't DOC land, have human activity commercial or residential, and is commonly land that is not pristine)
Number of Records	Composite sample, 6 spots in the creek, 10–30 species per sample, species list of up to 150 macroinvertebrates, 100 sites altogether some discontinued, some new. Currently 35 sites. Approximately 250 total samples.
Approach to Metadata Management	
Data Standards	**Rank:** 2 **Standards Used:** MCI **Notes:** Really more of an indicies, try to standardise with other regional councils, e.g. habitat assessments, level of resolution of speciation
Data Dictionaries	**Dictionaries used:** None really **Notes:** List of taxa names from Cawthron, Cawthron identifies to MCI level, we enter in to our main spreadsheet.
Data Management Policy / Procedures	No formal protocols
Data Quality	**Rank:** 3 **Notes:** No defined certainty statements. Collection protocols are defined and nationally agreed upon. Data are entered once and not double checked. Mistakes are only really around data entry
Future Development	**Proposed development path for the database:** No concrete plans, but would like to change the format in the future to make it more user friendly for analysis and state of the environment reporting **Extent the data are continuously being maintained:** Updated annually **Certainty of the future of the data:** Well assured

Development, Integration and Interoperability

The following are responses by West Coast Regional Council to questions on future biodiversity database developments and integration/interoperability issues.

New Databases Planned	
Hardware and Software Platforms	None.
Integration with Major Biodatabases	**Ability to provide or exchange very basic collection or observation data:** Excel, Hilltop **Other fields that might be of benefit to major biodatabases:** Data are public, can be given as a whole file as required.
Barriers to Integration with	Dissolved oxygen ph, tebidity, resuspendible solids, other physio chemical data.

Major Biodatabases	
Benefits of Integration with Major Biodatabases	To WCRC – have our own suite of environmental conditions here but there are similarities with certain land uses in other regions. So it could be useful to conduct comparisons with other regions. To others – again comparisons. Scientific benefit to researchers in Universities and CRIs. Relational capital from cooperating. Can see the big picture benefit.
Externally Stored Data	**Data currently stored in external organisation's systems:** NZFFD
Integration with other Regional Councils	**Benefits of data integration:** Comparison benefits **Benefits of sharing systems/infrastructure:** Centralised system – if there were better quality control components for example would give us a better result **Examples of unintentional or unavoidable reinvention of the wheel:** When I worked with NIWA Auckland on urban streams for ARC, looked at all cities for Auckland, characterised ecological communities. Another firm were doing the same thing at a smaller scale for North Shore City Council. No reinvention that I'm aware of in terms of systems. Certainly true in terms of approaches and techniques others are using and innovation in these. Only find out by going to conferences.