DISTRIBUTED INFORMATION ACCESS IN NEW ZEALAND

CLIENT:	FRST
PROJECT:	PGSF Contract UOO621 Distributed Information Systems
VERSION:	2-02
VERSION STATUS:	RELEASE
LAST UPDATE:	31 JANUARY 2001
AUTHORS:	HOWARD NICHOLLS (ALCHEMY GROUP) ROBERT GIBB (LANDCARE RESEARCH)
ID:	DIS-RPT-002

Alchemy Group Limited, PO Box 2386, Christchurch, NZ.
Ph +64-3-962-0396, Fax +64-3-962-0388, www.alchemy.co.nz

Manaaki Whenua Landcare Research, Private Bag 11 052, Palmerston North, NZ,
Ph+64-6-356-7154, Fax +64-3-355-9230, www.landcare.cri.nz

CONTENTS

1. DOCUMENT PURPOSE AND SCOPE

The purpose of this document is to describe the key technology issues for distributed information access in New Zealand. It is written from an industrial and public sector perspective, representing the views and findings of a wide cross-section of institutions in public and private sectors. It is an output of Objective 2 of the Distributed Information Systems project funded under contract UO0621 by the New Zealand Foundation for Research, Science and Technology (FRST).

It complements other project material produced by the academic research team at the University of Otago and its collaborators.

It focuses on requirements and applications, and is intended to provide a real-world, New Zealand-oriented context for the research in distributed information technologies (DIST).

The report represents the culmination of a series of workshops, industrial consultations, a questionnaire, and the experiences of the authors' institutions during the project, and therefore it supplements any previously produced material.

2. BACKGROUND TO NZ DIS PROJECT

2.1 Overall Drivers for the Research

Rapid developments in automated data acquisition, telecommunications, data storage capabilities, and Web-based interconnectivity have led to an explosive growth in available information resources. However, although these expanding information repositories are increasingly available to wider circles of society, they are archived using various media types (text, audio, images, digital media, etc.), stored in differing formats (flat files, relational databases, object-oriented databases, etc.) and organised according to differing semantics. The result has been that the expansive growth in the volume of information has outstripped the capacity to organise, access, and interpret efficiently the needed information from these vast stores.

The goal of the NZDIS project is the development of advanced distributed information system technology (DIST) that will enable New Zealand users and enterprises to access and integrate information from all available sources in a timely and useful fashion.

The strategic importance of such technology for the welfare and competitive pre-eminence of national economies is widely recognised and has led to the initiation of ambitious national programmes in the United States, Singapore, and elsewhere. In these cases there is a recognition that technological development must be complemented by programmes to promote widespread acceptance of the new technology and accelerate technology transfer.

The DIST programme is in line with these other initiatives and will carry out research and development appropriate for New Zealand's needs and will establish the foundation for the country's distributed information research.

2.2 Government and Public Sector Issues

During the decade before the NZDIS project, the distributed computing environment changed radically. In the late '80s, distributed computing was characterised by loosely networked computers using email, telnet and ftp over relatively slow inter-city and inter-organisational network links. Only large organisations with offices and staff spread across the country could afford the costs of participating in distributed computing. Large centralised computers served the needs of multiple government departments via networks of terminal servers, DSIR divisions ran a network of Digital VAX computers networked together using common software, and Universities were linked in a separate network.

These separately managed networks were connected at a few points where low volume traffic between organisations was guardedly permitted. The regulatory environment prohibited computers from one organisation using another organisation's network to link them. Data was more often transferred via tape than network — one government computer systems manager's attitude to networks was encapsulated by the comment that the bandwidth and cost effectiveness of a truck full of (40MB) tapes far outweighed the capacity or convenience of any network.

Amongst researchers, the primary concern of users of 'distributed information' was in defining data transfer formats, and ensuring they had software to read the (tape) format and convert data between formats. Some groups of researchers in related disciplines were tackling the concepts of common data standards, but often constraints inherent in a standard were seen to outweigh advantages. Inter-governmental needs were dominantly met through sharing a single very large computer running a shared suite of software — so data exchange was typically an issue of negotiating data access privileges on a single computer rather than negotiating network capacity or data transfer formats.

By the mid-'90s, the central shared government computing facilities had been largely dismantled and replaced by departmental networks of personal computers on high-speed LANs. Each government department, released from the shackles of a centrally controlled computing resource that only partially met its needs, and faced with a rapidly changing government environment, implemented its own computing system focusing dominantly on its internal needs. Inter-government agency network links were often still limited in scope.

In contrast, the research community — which had always been at the forefront of the use of inter-agency networks — took the opportunity of a liberalised regulatory environment and rapidly decreasing costs of network capacity to combine their networks into a single shared wide area network.

Deregulation of the telecom industry led to telecom providers offering enhanced data services, including Internet services, as part of a network connection. These changes were enormously liberating. For the first time organisations in different cities could form a loose network merely by each leasing a connection to their local Internet service provider (ISP) and then arranging for access to each other's services. Previously they had to lease lines between each of their sites and agree to share the not insignificant costs. Coupled with substantial increases in network bandwidth, this opened the door to the possibility of data and service provision and sharing on an unprecedented scale.

These fundamental changes came with a raft of new issues. The old issues of data transfer standards and formats were now joined by new technical issues and possible solutions such as: data description (meta-data); data discovery (data directories, robots); data presentation (mark-up languages); enhanced security issues; collaboration; collaboration protocols; dynamic bandwidth negotiation (ATM); network languages (java) …

Further, the new capability raised user expectations and whole new industries based on different combinations of these issues and new business models have grown round the new services — in short, a digital revolution and the information age.

This was the environment at the inception of the NZDIS project.

2.3 Industry Issues

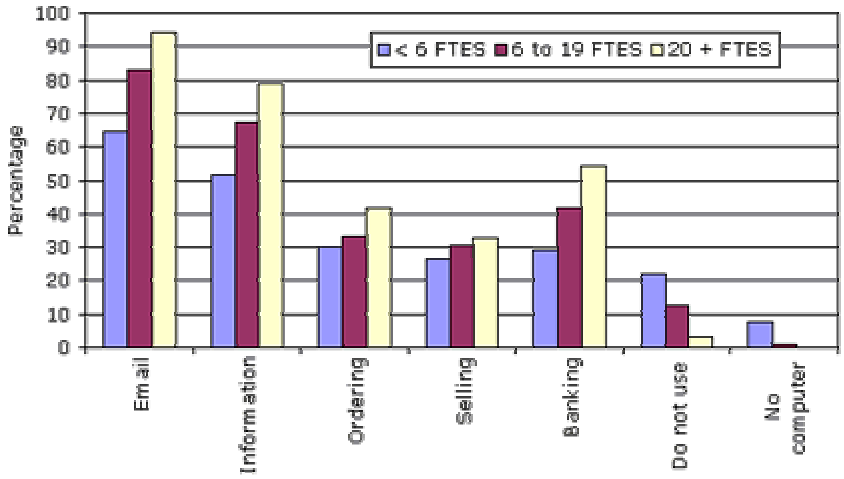
Many of the issues encountered by the public sector were mirrored in the private sector. A key difference is that in the private sector, most organisations are much smaller and have a very focused set of functions required from their information systems. As with the public sector, Internet access has transformed the opportunities for small organisations with low-end desktop technology to search and acquire information from a diversity of sources.

However, current use is generally at the level of site searching, not the execution of queries nor combing information from multiple data sources (see charts, below). Use is generally in passive mode, with input of data and e-commerce not yet fully taken off.

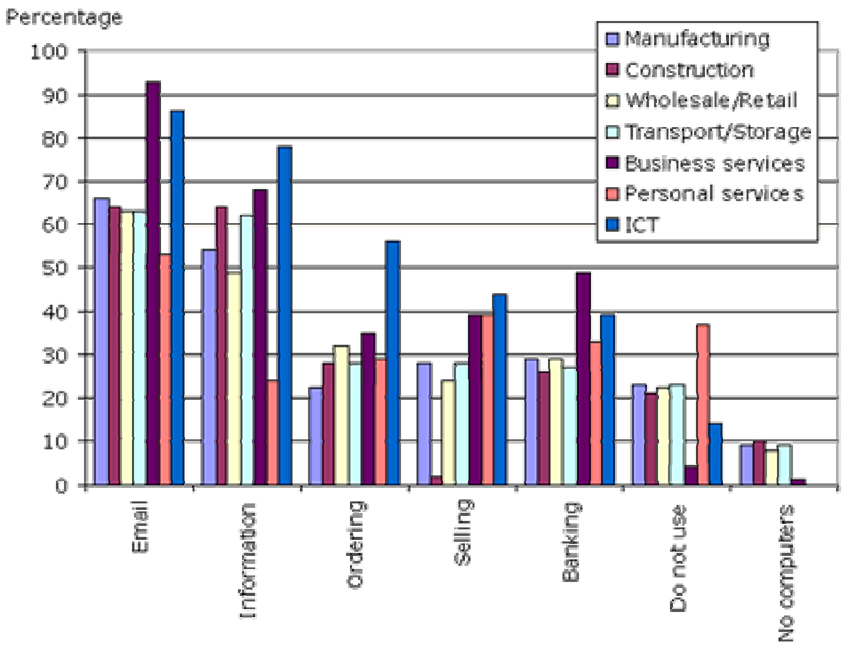
A survey of New Zealand businesses was commissioned by the Ministry of Economic Development to provide background information for the Government's ecommerce strategy. BRC Marketing and Social Research conducted a telephone survey of 506 respondents in August 2000. The following results extracted from the report (www.med.govt.nz/irdev/elcom/survey/) add weight to the contention that modern businesses have very strong information needs and view investigative search as a vital part of their businesses.

Use made of the Internet can be roughly graduated from casual through to sophisticated, with simple email use at one extreme and the use of interactive and secure transactional web sites at the other. For each use of the Internet there is a consistent correlation with the size of the firm. Email is the most used service and is taken here as being a useful surrogate measure of internet connection. Next in importance is information gathering followed by ordering, banking and selling. It appears that medium (6 to 19 FTES) and larger firms (20 or more FTES) use the Internet for banking more than for ordering or selling goods and services.

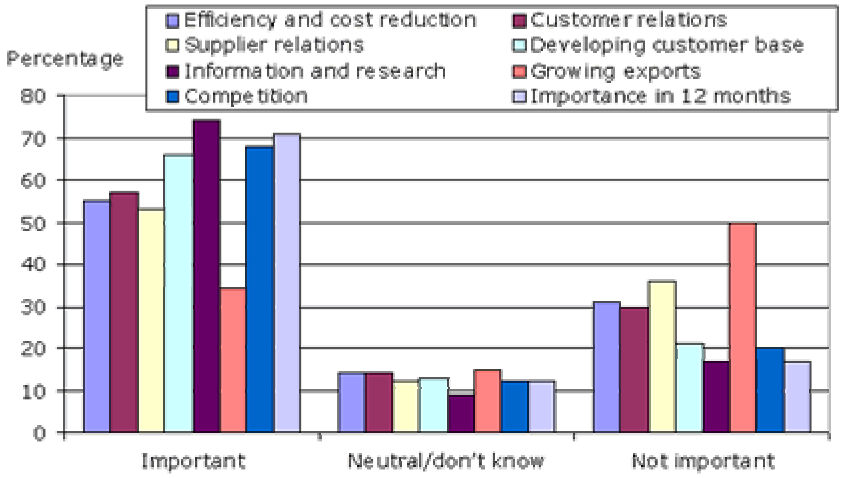
Almost all larger firms (94%) are using the Internet for email now and medium sized firms are approaching this level (83%) however only 65% of small firms (fewer than 6 FTES) are currently using email. Of the identified Internet uses, selling of goods and services to customers appears to be the least important function.



As seen in the following graph, analysis of use of the Internet according to business sector gives broadly similar results to those above, in each category. Following use of email, the most important use of the Internet is in information gathering.



The survey also attempted to gauge business perceptions of how important e-commerce was considered to be, and for what reason. The use of the Internet for information and research is seen as the single most important factor.



Given this overarching context, the specific needs and issues of distributed information access are discussed in the next section, in particular through a survey conducted by the NZDIS project and via a series of workshops and industry consultations.

3. REQUIREMENTS FOR DISTRIBUTED INFORMATION SYSTEMS

3.1 Overall Requirements

To move from simple querying of single datasets to more complex, multi-source querying, the requirements are demanding and the technologies still in an emergent stage.

Derived from the issues identified above, the key requirements for distributed information access include the need for meta-data standards, more flexible environments for integrating heterogeneous data sets, and greater attention to security and access control. More specifically, the requirements are:

Overall:

- to be able to integrate systems with differing meta-data;
- to be able to integrate systems with differing storage formats and media;
- to be able to incorporate legacy systems and unstructured datasets;
- customisable, to suit different application domains.

More specifically:

- have a communications architecture for passing messages, typically in a dialogue, between the distributed components of the system;
- have a means of describing models of the user (application) domain and of data sources;
- have a means of storing and accessing these models;
- have a means for expressing queries such that:
  - a single query must be able to refer to information in multiple data sources, and
  - the query must be able to be broken into sub-queries but the results combined into a single result set for presentation back to the user;
- have a means of accepting queries and delivering results back to the user in a variety of media;
- have a secure means of access control, and protection of data integrity from unauthorised actors in the distributed environment;
- have an open architecture that can accommodate new developments in distributed software engineering technologies.

In an Internet context where such facilities are available to the general public, there are additional performance issues introduced:

- the number of users may fluctuate and be unpredictable;
- the concept of a stateful user session is harder to maintain in large concurrent user bases;
- the interacting parties may belong to independent organisations that do not necessarily trust each other;

- the communication infrastructure does not provide quality of service guarantees.

3.2 DIS Technology and Tool Requirements

The overall requirements driven by private and public sector needs led to the identification of requirements for, and development of, the following distributed information system techniques and tools. These techniques and tools are comprehensively documented, and further information can be sourced at the NZDIS website, at http://nzdis.otago.co.nz . They were produced by the NZDIS research team.

Specific required technologies and tools were:

- The NZDIS framework – a system for integrated querying of distributed heterogeneous data sets which required several key components: (i) user interface agent, (ii) query processing agent, (iii) executor agent, (iv) data source agents.
- The NZDIS agent platform – underpinning the NZDIS framework is the concept of software agents. To facilitate the rapid development of software agents, it was found that a platform for simplifying their development was required, to be made available as a set of Java APIs. The platform permits the construction of logic-based messaging agents using object oriented programming techniques.
- Knowledge Interchange Format (KIF) parser and pattern matcher - the Knowledge Interchange Format (KIF) is a language of first order logic used in the agent and expert systems community for knowledge interchange. To use KIF effectively, it was required to produce a package for Java programmers which has full parsing and pattern matching capabilities for ANSI KIF format.
- OQL 3 parser and query analyser - the Object Query Language (OQL) is a standard language (http://www.odmg.org) for querying object-oriented databases and object-oriented data. The project required a parser to perform full parsing capabilities for OQL 3.
- A set of Java classes and utilities for NZDIS project were required including an event manipulation framework, an observable object store, command line processing, and others.
- A Petri Net tool - a Java-based, object-oriented Petri Net simulation framework was required.
- A Java-based general-purpose visualization module was required for the Petri Net (JFern) tool and was made available for other projects. This graphical toolkit was based on JHotDraw package.
- A lightweight, generic, highly customisable and efficient graphics package was needed for drawing and manipulating diagrams containing nodes and links. It also needed features to allow automatic layout. The package implemented, named Swift, was written in Java, and will work with either the Java 1.1 or 1.2 graphics APIs. Swift was required in the NZDIS system for drawing agent interaction diagrams.
- An OQL planner (Java version) was needed - the Object Query Language (OQL) is a complex language that specifies what needs to be done but not how to do it. The OQL planner is a package that takes an OQL query and breaks it down into a set of simple tasks that may be performed efficiently by an executor. It also performs some ontology translation on the query to turn it into a suitable form for planning. Output is in XML.

- An OQL planner (Mercury Version) was also needed - the Mercury OQL planner is another planner that accepts OQL as input and produces XML as output, but is designed using different technology to make it more powerful and extensible. The system implemented is written in Mercury (http://www.cs.mu.oz.au/research/mercury/), and converts the query into a mathematical form known as a monoid comprehension. It currently relies on the ontology translation mechanism from the Java planner.

The following tools were required for use with the Mercury planner, but as separate packages that may be used alone. Most of these tools and further information are available on request from mpurvis@infoscience.otago.ac.nz.

- Mercury pretty-printer provides which provides tools to easily format output of Mercury programs. It uses indenting to display structure.
- Mercury XML parser/printer. The NZDIS system uses XML to read and store complex information. This package can be used to easily parse or print the most commonly used subset of XML and allows Mercury programs to input or output complex information in a consistent format.
- Mercury XMI parser. XMI is an XML vocabulary that can be used to store the information from a UML diagram. The XMI parser takes the output from ArgoUML (A UML editor) and converts it into an internal data structure.
- Mercury OQL (Object Query Language) parser/printer. This is a tool that parses and prints OQL queries for Mercury.

4. TESTBENCH: PROCESS MODEL FOR A DISTRIBUTED QUERY

A distributed query application was defined for use as a test bench for the system being developed in Objective 1. This application is derived from a real-world scenario and requires integration with multiple commercial datasets. It was developed as part of Objective 2 of the research programme. The datasets were sourced by and through the industrial connections of the Objective 2 subcontractors.

This query is documented through a traditional Process Model published on the web at http://nzdis.otago.ac.nz/schematic/nzdis.html. Elements of the model are shown below to illustrate the query.

The query, which is in the Real Estate domain, can be phrased as follows:

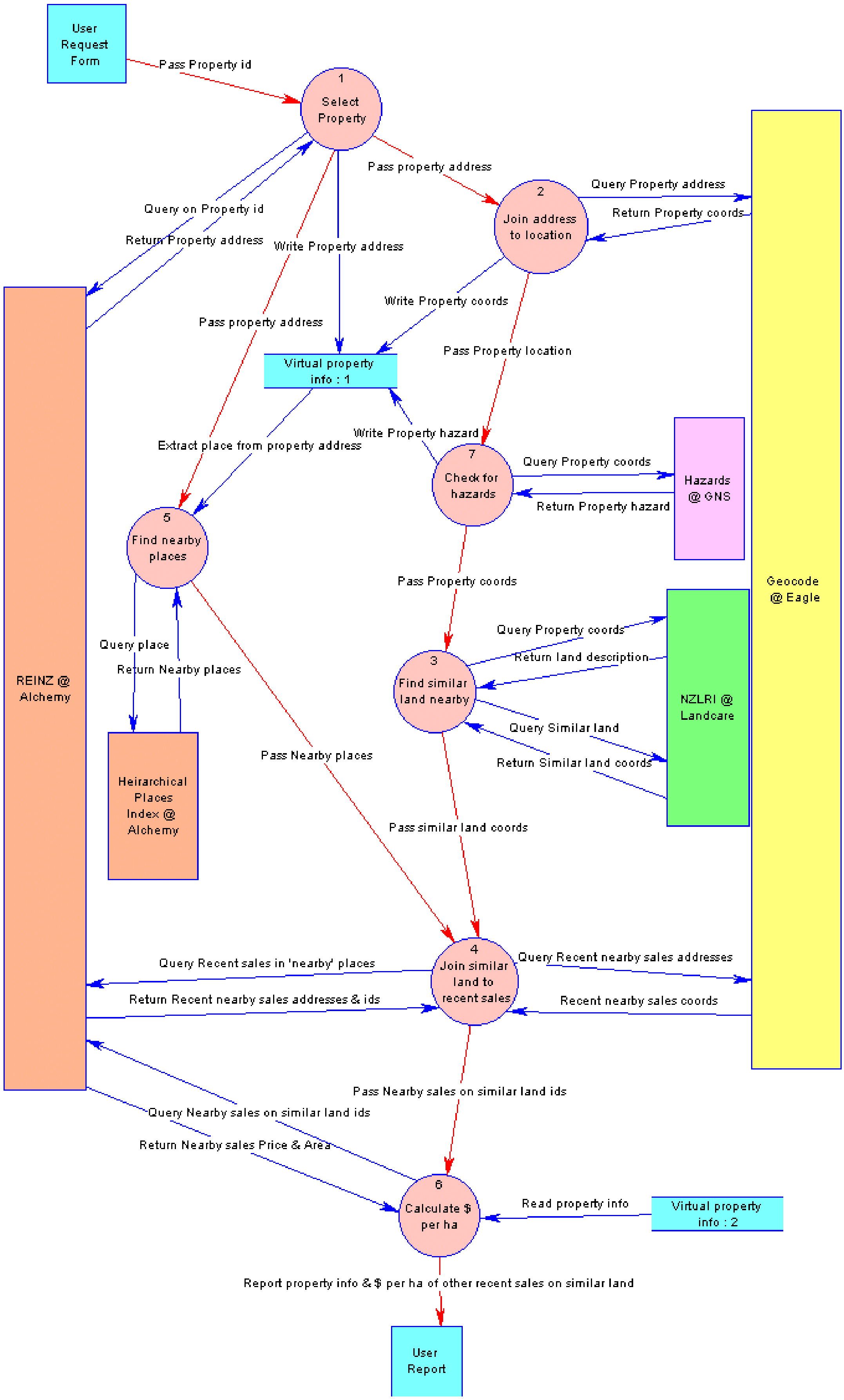
For a given (rural) property show any natural hazards (documented by GNS), the land-use capability (as assessed by Landcare Research), and the value of nearby properties of a similar capability and hazard that are reported as recently sold (in the REINZ database).

Eagle Technology's Geocode application may be used to join address style locations with map grid coordinate style locations if required.

The Process Model serves both to document the query and to illustrate a possible solution — though the query optimiser need not come up with the same solution. The reference to GeoCode is provided but in a real application it is not anticipated the user would have to know that it were required; similarly, the explicit references to other data sources would not be required for users to provide up front, but would be offered by the system as part of possible solutions for a user to confirm.

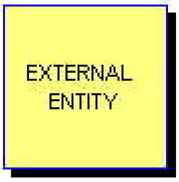
The process model is shown overleaf. The components of the query are as follows:

Entities (rectangles)	Data Flows (arrows)	Processes (circles)
1. User Request	1. Query Property Id for Address	1. Select Property
2. Real Estate Institute of NZ (REINZ)	2. Write Property Address	2. Join Address to Location
3. Hierarchical Place Index	3. Query Property Address for Coords	3. Find Similar Land Nearby
4. Hazards	4. Write Property Coords	4. Join Similar Land to Recent Sales
5. NZ Land Resource Inventory (NZLRI)	5. Query Property Coords for Hazard	5. Find Nearby Places
6. Geocode	6. Write Property Hazard	6. Calculate $ / Ha
7. Property Info	7. Query Property Coords for Land Description	7. Check for Hazard
8. User Report	8. Query Similar Land for Coords	
	9. Query Place for Nearby Places	
	10. Query Recent Sales for Address Ids	
	11. Query Nearby Sales for Coords	
	12. Query Nearby Places for Area Sale Price	


User Request Form
Pass Property id
1 Select Property
Pass property address
2 Join address to location
Query Property address
Return Property coords
Query on Property id
Return Property address
Write Property address
Write Property coords
Pass property address
Pass Property location
Virtual property info : 1
Write Property hazard
Extract place from property address
7 Check for hazards
Query Property coords
Return Property hazard
Hazards @ GNS
5 Find nearby places
Pass Property coords
Geocode @ Eagle
Query Property coords
Return land description
3 Find similar land nearby
Query place
Return Nearby places
NZLRI @ Landcare
REINZ @ Alchemy
Query Similar land
Return Similar land coords
Pass Nearby places
Heirarchical Places Index @ Alchemy
Pass similar land coords
Query Recent sales in 'nearby' places
4 Join similar land to recent sales
Query Recent nearby sales addresses
Return Recent nearby sales addresses & ids
Recent nearby sales coords
Pass Nearby sales on similar land ids
Query Nearby sales on similar land ids
Return Nearby sales Price & Area
6 Calculate $ per ha
Read property info
Virtual property info : 2
Report property info & $ per ha of other recent sales on similar land
User Report

An example each of an entity, process and data flow follows:

i) Example Entity: 6 - GeoCode

	**Geocode**
Description:	GeoCode is an on-line application providing an automatic translation from street addresses within New Zealand to NZMG coordinates.
Institution:	Eagle Technology
Administrator:	Peng Aik Lim
Fields:	Input a DBase file containing `Address, Fine` and `Coarse,` with `Fine` and `Course` taken from places.txt and receive an ESRI shape file back containing matching NZMG coords
Statistics:	
Dataset:	Send DBase file as an e-mail attachment to mailto://geocoder@eagle.co.nz/

ii) Example Process: 3 – Find Similar Land Nearby:

	**Find similar land nearby**
Description:	This procedure needs to find a sufficiently large list of nearby areas that have similar land to get a non-null set when joined with land that has sold recently. A strategy is proposed whereby the NZLRI is repeatedly queried for the number of map units with similar Land Use Capability (LUC) starting with a very tight LUC constraint and slackening the constraint until a target number of map units is reached. The target number of map units will have to be learnt from experience with the REINZ database. Once the target number is reached, the same constraint is used to extract the coordinates for the set of map units.
Comments:	**Similar land** is interpreted in terms of LUC, so the first step is to query the NZLRI database for the LUC at the given coords. LUC is returned as a triple of *class, subclass* and *unit*. **Very similar** land has the same *class, subclass* and *unit*. **Similar** land has matching *class* and *subclass,* while **slightly similar** land only matches *class*. **Nearby** is interpreted as within the borders of the Regional Council with the same name as the *District* field in REINZ. If Regional Council boundaries are found to be too loose a constraint, either a circle of varying radius or District Council boundaries could be used to constrain the search.
Code URL:	

iii) Example Data Flow: 7 – Query Property Coords for Land Description

Query Property Coords for Land Description

Description: Given a coord pair, return the Land-Use Capability Class, subclass, unit and area (in Ha) for the map unit surrounding the point. An optional search radius may be provided, in which case the information will be provided for all map units that overlap a circle of that search radius. The areas provided are the areas of the map unit, not the area of intersection of the map unit with the circle.

Institution: Landcare Research

Author: Robert Gibb

Query: Arc/Info AML pseudo-code fragment

```
ArcPlot: &args x y r
ArcPlot: /* x is NZMG Easting in m
ArcPlot: /* y is NZMG Northing in m
ArcPlot: /* r is optional search radius in m
         /* (if provided the circle reselect param
         /* will be used)
ArcPlot: mape /data/base/gis/nzlri/silri-idx
ArcPlot: reselect /data/base/gis/nzlri/silri-idx poly ~
           { one %x% %y% | circle %x% %y% %r% passthru }
ArcPlot: &type Class,Subclass,Unit,Area
ArcPlot: &do i = 1 &to [show select ~
           /data/base/gis/nzlri/silri-idx poly]
ArcPlot:  &sv luc = [show select ~
            /data/base/gis/nzlri/silri-idx poly %i% ~
            item luc]
ArcPlot:  &sv ha = [show select ~
            /data/base/gis/nzlri/silri-idx poly %i% ~
            item areah]
ArcPlot:  &type [substr %luc% 1 1],[substr %luc% 2 1], ~
            [substr %luc% 3 3],%ha%
ArcPlot: &end
ArcPlot: &type eof
```

Response: Comma separated ascii data comprising:

```
Line 1    : Class,Subclass,Unit,Area
Line 2    : {1..8},{c,w,s,e},{1..99},{1..999}
                                  ...
Line n    : eof
```

5. INDUSTRIAL CONSULTATION AND EXAMPLES

5.1 Industry Consultation

A series of demonstrations, workshops and seminars were held in relation to the project.

Public Workshops and Seminars:

- NZ Distributed Computing '98 @ Te Papa (NZDC'98), R. Gibb, M. Purvis, N. Ward, and A. Bond, http://nzdis.otago.ac.nz/servlet/Dis/nzdc98.html, Wellington, September 3-4, 1998.
- Officials Information Management Meta-data Working Party workshops, Landcare Research, Gibb, R., Palmerston North, March 5th, March 26th and 30th April 1999. (For details see http://www.massey.landcare.cri.nz/~rgg/mfe/oim/ and Appendix A: "Land Information New Zealand's Contribution to the Development of a State Sector wide Standard for Meta-data" in *Report of the E-Govt Meta Standards Working Group*, http://www.govt.nz/egovt/mdwsg.pdf).
- "Tools for E-Commerce", Industry seminar, H. Nicholls and D. Ashby, Christchurch, 17 August 2000. Attended by over 50 participants from the software industry and local government.
- NZ Distributed Computing 2000 (NZDC'2000), M. Purvis, W. Sandle, and participation of Telstra, Sybase, and DSTC, Auckland University of Technology, http://nzdis.otago.ac.nz/nzdc2000/index.xml, Auckland, November 9-10, 2000. The NZDIS prototype system was demonstrated at this workshop.

Demonstrations:

- "Distributed Information Systems", M. Purvis, G. Bush, R. Ward, M. Nowostawski, and D. Carter, Computer and Information Science Seminar Series, University of Otago, September 8, 2000.
- "The New Zealand Distributed Information Systems Project", M. Purvis, Auckland University of Technology, October 25, 2000.
- "The New Zealand Distributed Information Systems Project", M. Purvis, M. Nowostawski, G. Bush, and D. Carter, *Complex Systems 2000 Conference*, Dunedin, New Zealand, November 19, 2000.
- "The New Zealand Distributed Information Systems Project", M. Purvis, G. Bush, *12th Annual Colloquium of the Spatial Information Research Centre (SIRC'2000)*, Dunedin, New Zealand, December 12, 2000.

5.2 Survey

A questionnaire was produced to canvass the views of industry on their needs for distributed information access. The questionnaire was sent to all 140 listed companies on the NZ stock exchange. The questions and their responses are shown below. As can be seen, the results are not particularly conclusive, so 1-1 consultations were held along with workshops.

Distributed Information Systems Questionnaire & Results

Thirty-five of 140 questionnaires (25%) were returned. Not all questions were completed on all forms.

1. Briefly describe what your business does

- see question 5

2. Briefly describe your role in the company

- CEO, CIO, Company Secretary, IT Manager

3. When you think of information for your business, what do you immediately think of and why is that information important to you?

4. Briefly describe your companies risk averseness with respect to implementing leading edge information technologies

- overall respondents were predominantly risk averse, but will use IT if benefits can be shown and provides competitive advantage

5. Which of the following best describes your business sector? (please tick only one)

Primary industry	10%	Finance, insurance	8%
Electricity, gas, water supply services		Manufacturing	20%
Construction		Information technology	3%
Wholesale	3%	Property	6%
Retail	3%	Education	
Tourism	14%	Health	3%
Transport, storage	6%	Cultural	
Communications	6%	Personal	
Research and Development	3%	Other (Hospitality)	3%

6. Which of the following best describes your business operation? (tick all that apply)

Local	11%
National	32%
International	57%

7. We wish to understand how the concept of meta-data is understood in industry and business. Is this term in active use in your organisation?

Yes	6%	No	94%

Please indicate the extent to which you agree or disagree with the following statements

		Strongly Disagree			Strongly Agree
8	**The WWW plays a minor part in our business's marketing strategy**	8%	30%	25%	37%
9	**The Internet can be used to keep in touch with our suppliers**		18%	40%	37%
10	**Management can use the Internet effectively to communicate with staff**	11%	34%	17%	28%
11	**Staff can use the Internet effectively to access business information**	11%	11%	34%	34%
12	**Our suppliers make more use of the Internet than we do**		54%	23%	11%
13	**A small proportion of our clients have access to the Internet**	6%	46%	31%	8%
14	**Obtaining reliable figures on Internet use by our suppliers and clients if difficult**	3%	20%	43%	28%

15	**As an operational tool the Internet is too expensive to be worth looking at**	51%	37%	3%	3%
16	**IT is essential to the operation of our business**		8%	17%	68%
17	**The availability of outside data sources has a significant impact on our strategic planning**		20%	60%	17%
18	**The integration of multiple sources of information is a significant issue in our organisation, and we are seeking improvements in this area**	8%	11%	57%	11%
19	**Our business has ready access to all significant electronic data and information sources**	14%	46%	26%	11%
20	**Our business experiences no difficulty in obtaining electronic data and information from other agencies**	11%	40%	31%	14%
21	**Our business will increase its use of electronic information in strategic planning**	3%	14%	60%	20%
22	**Data security is an important IT issue at our company and significantly affects strategic planning**	6%	26%	46%	20%
23	**Charging for access to data and information will become an increasingly important issue in our organisation**	26%	31%	23%	14%
24	**Delivery timeframes for data and information have a significant impact in our organisation**	3%	11%	51%	31%
25	**Our business considers data quality descriptors must be provided with data and information**	3%	43%	34%	14%
26	**Electronic data and information supplied to our organisation is generally of high quality**	6%	31%	49%	11%
27	**Government agencies are seen in our organisation as key data suppliers**	23%	31%	37%	6%

5.3 Organisations Consulted

In addition to the questionnaire, the following organisations were consulted in the course of this research, including (in most cases) via their participation in various workshops.

Organisation type + sector	Name
Commercial companies	16 companies
CRIs	Forest Research
	Industrial Research Limited
	Land Information NZ
SOEs	Asure NZ
	Airways Corporation
Government Organisations/Agencies	Antarctica NZ
	Dept of Conservation
	Dept of Social Welfare
	GOVIS
	Hillary Commission
	Land Information NZ
	Min for Environment
	Min of Agriculture & Fisheries
	Min of Commerce
	Min of Justice
	Min of Research, Science & Technology
	Min of Transport
	National Archives
	National Library
	NZ Immigration Service
	Southland RC
	State Services Commission
	Statistics NZ
	Treasury

5.4 Sample Applications

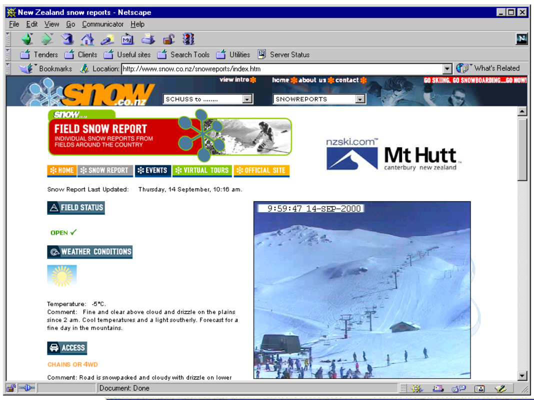
The following are applications developed by the project team, not directly as part of the funded project, but using the experiences of the DIS research to influence their design. They illustrate the wide range of inputs and outputs that can eventuate in multi-party information systems. The examples complement the multiple querying information access supported directly by the research.

5.4.1 www.snow.co.nz

www.snow.co.nz is New Zealand's leading portal for the skiing industry. It is *the* site to go for information on latest snow reports, ski field facilities, events details, employment, equipment and the latest avalanche situation. Its business model was developed by Internet Management Solutions Ltd and its underlying database, input system and delivery mechanisms were developed by Alchemy Group. It has evolved into a site handling over 1 million user sessions per year.

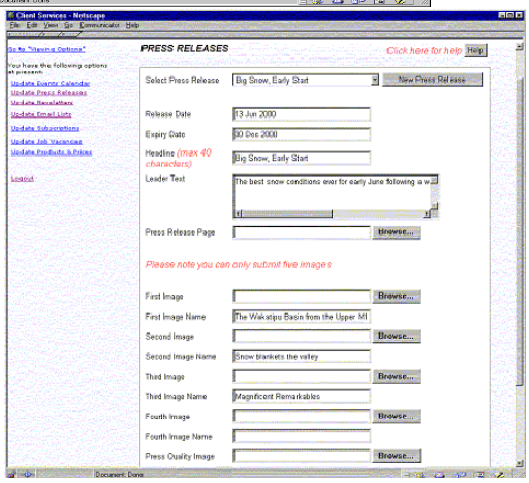
Its core functions revolve around distributed input of a diversity of data, including:

- snow reports filed by staff online at the skifields;
- hourly updates from webcams on the skifields;
- photographs of the fields submitted by staff online at the skifields;
- media releases and daily newsletters input via an online content management interface;
- 360° panorama imagery for virtual tour display;
- events calendar details of events, dates, locations etc.;
- employment opportunities filed by employers online to the content management interface.



Its outputs are also diverse, covering a variety of media and presentation formats:

- web browsers;
- Reuters news feed;
- GSM short message service text messages;
- email;
- fax delivery.

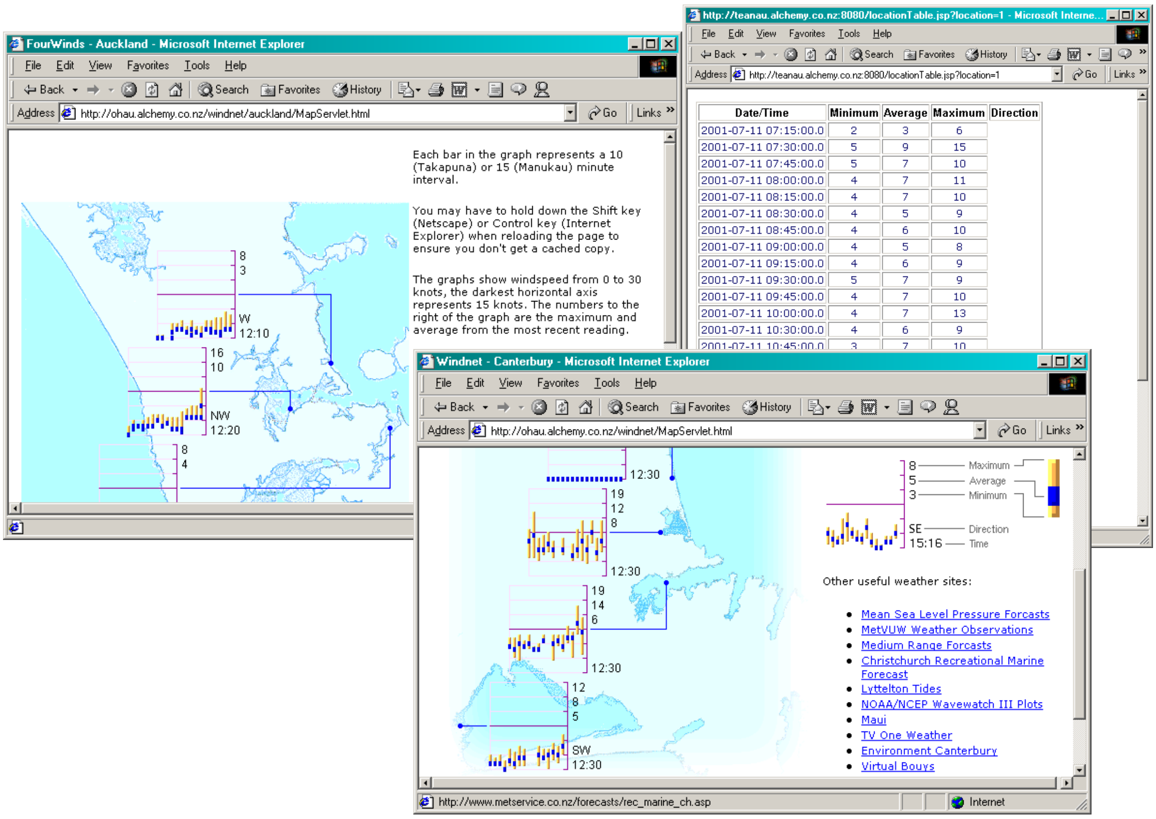


The integration of current ski field photos into the site involved remotely uploading the digital image from a camera at each ski field and automatically processing it for display on the web page. As a result, each major ski field has an up-to-date picture on the site, taken each hour during the morning and also in the late afternoon.

Interface modules were developed to distribute the snow reports from the www.snow.co.nz database to other organisations and news providers. These include Reuters, who receive a summary of the snow reports for their own distribution, and Vodafone GSM Short Message Service, who distribute summary snow reports to the subscribers' cellphones.

5.4.2 www.windnet.co.nz

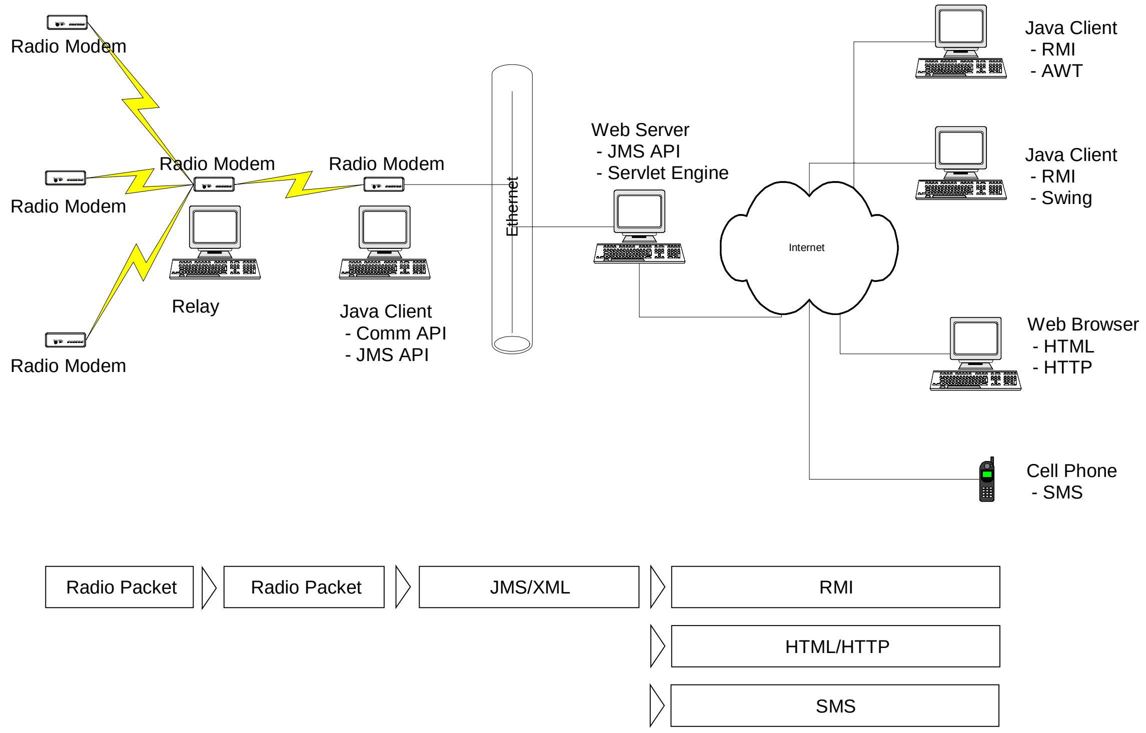
Experience on the DIS project has been used in developing a highly distributed application which manages real time data from remote sites via radio and disseminates it to users via various media including web browsers, text-based cellphone messaging, pagers and faxes (www.windnet.co.nz).



The system was based on Java 2 technologies and was heavily distributed. Radio modems at locations around the coast sent packets to a relay station which then forwarded them to a base station. The base station then used Java Messaging Service and Comm APIs to communicate data to the Windnet server. The server can receive data from multiple base stations via JMS messages.

The server used a Java 2 servlet engine to deliver formatted output data from the data received from the base station to a variety of clients. The clients included those running a custom client user interface built with the Java AWT API, those running a custom client user interface built with the Java Swing GUI API, standard web browsers receiving HTML via HTTP, and cellphones receiving SMS text messages.

This highly distributed information system is illustrated overleaf. A key feature of the system software was the separation of content from message protocol and from presentation format, maximising the versatility of the system.


Radio Modem
Radio Modem
Radio Modem
Radio Modem
Relay
Radio Modem
Java Client
- Comm API
- JMS API
Ethernet
Web Server
- JMS API
- Servlet Engine
Internet
Java Client
- RMI
- AWT
Java Client
- RMI
- Swing
Web Browser
- HTML
- HTTP
Cell Phone
- SMS
Radio Packet
Radio Packet
JMS/XML
RMI
HTML/HTTP
SMS

6. E-GOVERNMENT INITIATIVE AND DIS

At the outset of the NZDIS project in the mid- to late '90s, government IT departments that had been focused on isolationist provision of internal departmental computing needs found themselves facing ever increasing demands for efficiencies. One area of continuing concern was to eliminate duplication between agencies by replacing them with shared data and eventually network-based systems.

Another driver for change was increasing demands of open government, whereby members of the public expect government information to be made available to them with minimal fuss or delay – one initial strategy for achieving this was for each agency to establish a web site containing as much information as possible in the hope that most requests would be satisfied directly from the website.

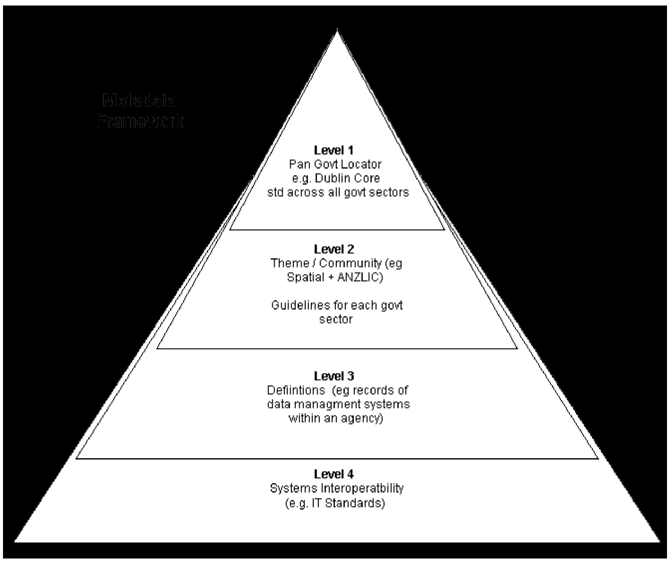
However, that raised its own problems in that members of the public had to be sufficiently savvy to find the right page on the right website. Without a common approach to web site structure and contents this was often a fraught exercise. Possible solutions included a common website or portal for all government information and a common directory or index. Such ideas depended on the existence of meta-data describing all Government information sources.

(See http://www.e-government.govt.nz).

6.1 Emergent Meta-data Standards

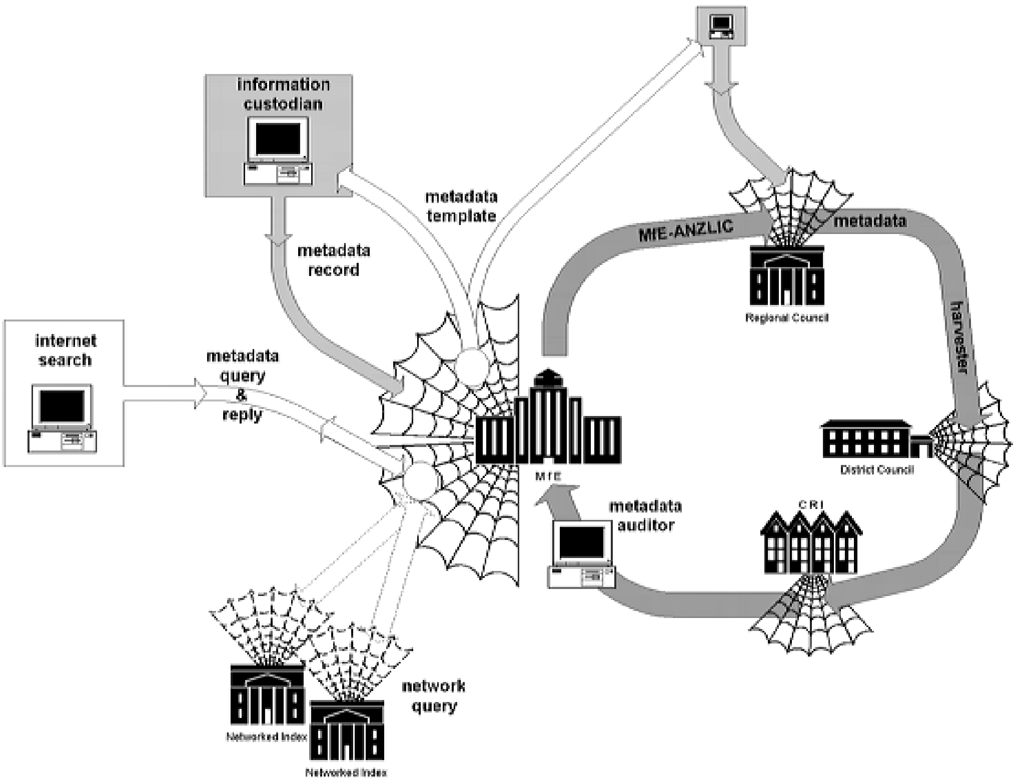
The Officials Working Group on Meta-data set out to raise the awareness of meta-data amongst key information managers within government agencies. Through a set of three workshops facilitated by one of this report's authors (Gibb), participants developed a common understanding of the role of meta-data in information management by which departments could assess both their common and unique meta-data needs, see how these related to each other and develop common approaches where appropriate. One outcome of these workshops has been a commitment across all government departments to the implementation of a common meta-data standard for government information discovery.

In an ideal world meta-data would be structured as an integral part of every data source. But that presents a dilemma — meta-data has many roles, ranging from discovery through management to representation and controlling usage. Some of these roles can be met by having meta-data embedded in the original document, which would imply a structure tailored to the needs of the software that is associated with the document type (text, spreadsheet, database, image, map, etc.), but other roles, such as discovery, are best met by having the meta-data available for searching in a central repository.



Historically this has been resolved through the development and use of a multiplicity of meta-data standards each focused on a particular suite of needs, such as

DublinCore, GILS, FGDC, ANZLIC. While computer systems were essentially independent entities, and networks were slow, this made a great deal of sense. However, with the globalisation of networks, people became aware that many meta-data standards had fields whose content was either identical or essentially equivalent. Acknowledging this trait, the current trend is to maintain meta-data in one form, typically that of the most demanding application that uses the meta-data, and then set up an automatic mapping of fields from the master system to the fields of each of the other meta-data systems to which the data needs to be accessible. The mapping can either be live, so that each in-coming query is translated into a query to the master system, or batch so that data is regularly reformatted and exported from the master system to other systems — typically at Level 1. This works because the requirement for currency is typically least stringent for Level 1 users.



Some new systems have sufficient inbuilt flexibility that these mappings are built into their design, which is often the case for emerging systems built round the XML standard. These concepts are illustrated by the diagrams taken from the work of the Meta-Data Officials Working Group. The first diagram illustrates the relationship between the four different levels of meta-data identified by the working-group. The second illustrates the flow of meta-data through a distributed system. Meta-data is collected and stored for its most demanding use (Level 3 — data management). Sufficient data to satisfy Level 1 use is regularly extracted by a robot to be placed, for efficient use for data discovery, in a central repository.

6.2 Government

Such a system was proposed for MfE and Regional Councils as part of the Environmental Performance Indicators programme before the e-Govt process got under-way. At the time it was recognised that it was inappropriate for MfE to progress its meta-data standards ahead of the rest of government, especially when a number of departments thought of meta-data in different contexts. This was one of the drivers behind the establishment of the Officials Working Group on Meta-data in early 1999. That has now evolved — separately from NZDIS — into a full study of metadata options for data discovery (Level 1) that will almost certainly result in a NZ version of AGLS (the Australian Government Locator Standard), which is itself a derivative of GILS (the US Government Information Locator Standard). A key difference between these emerging meta-data standards and their predecessors is that an assumption is made of a relationship between each field in the new standard and equivalent fields in other standards that exist at both the same level and lower levels, and that these relationships form part of the definition of the standard. Historically such relationships only evolved as cross-walk tables after the establishment of the standards.

6.3 Landcare Research

The concept of modular, interrelated meta-data standards has been developed further for Landcare's own (Level 3) meta-data standard. Landcare has a number of specialist scientific collections. Standards have evolved overseas specific to each type of collection (maps, geographic data, herbaria, insect and fungal collections), but Landcare needed an overarching set of standards that could both call upon the specific requirements of each discipline, and capitalise on commonalities where they existed. The parallels between these requirements and the requirements of a modern meta-data system were apparent, and the result is a modular system with numerous modules that can be called into play in any mixture for each specialist collection or its associated database. The fields within each module have explicit relationships with fields in other standard systems. At present the Landcare Meta-data Standard remains in draft form, but it was circulated amongst other agencies with overlapping requirements (such as DOC, MfE) and received favourable comment in mid-2000. Developments both in affordable off-the-shelf applications to suit such meta-data and structural changes in the way science data is to be managed within Landcare mean the probability of the meta-data standard being implemented in the near future.

6.4 NZDIS Meta-data Requirements

At its core, NZDIS is tackling issues of Level 4 meta-data. In contrast, the demand for meta-data in the New Zealand community is currently being addressed at Levels 1 and 2, with some organisations such as Landcare addressing Level 3 issues. The disparity between the current meta-data needs and experiences of the general NZ community and the NZDIS project means that while the project can draw on the experiences of the community for broad issues and context, there is little direct detail experience that the project can draw on. As has been demonstrated in this report, there are, however, many ways in which the project can contribute to the NZ meta-data community.

7. FUTURE ISSUES

7.1 Security

As organisations increasingly rely on distributed technologies, security and privacy become major issues requiring resolution. Increasing numbers of businesses are creating extranets, and the opportunity to integrate such systems further using DIS technology for complex multi-source searching adds another dimension of complexity. As more organisations move to such publicly accessible environments, malicious attacks against networks are increasing in frequency and sophistication. General public users have high concerns about privacy of their personal data, and businesses express caution at the risks involved in being part of a distributed environment.

Larger and technologically savvy businesses manage risk with various security solutions that monitor networks for vulnerabilities and threats, detect intrusions, and have policies to respond to attacks.

However, in a more strongly distributed world requiring multiple organisations to deliver components of a result to a query, each organisation becomes just part of the bigger security equation. There could be many networks or sub-networks, geographically and organisationally dispersed. Security policies across the organisational network may be inconsistent or not fully enforced, and opportunities missed to reduce risk.

Within an organisation, tools exist to scan for vulnerabilities, detect intrusions, and respond. They log significant volumes of information, which require specialist knowledge to collect, correlate and analyse. The logged security information itself comes from various sources within a large organisation. Therefore, DIS technologies themselves may assist the reduction of security risks created by distribution, by enabling better analysis.

7.2 Promotion of E-Government

The development of a NZ version of AGLS and its supporting infrastructure is an essential first step in developing a culture of information sharing. If we are to capitalise on these developments a great deal of work needs to be done, not only in government circles but also in agencies that have close links with central government, such as CRIs and local government. The DISABC project will have a significant role both in facilitating these developments and in developing tools that can be used in the wider community.

7.3 DIS and Applications for Broadband Communications

The project is continuing in 2000-2002, and will build on the results of the current project to carry out the following tasks:

- Develop an enhanced query planning system that utilises collected system performance information for the optimisation of queries in a distributed environment.
- Develop a suite of DIS agent conversation protocols and their implementations for system agent message passing in the DIS framework. The NZDIS agent-based architecture supports the use of agent conversation protocols for flexible, extended information interchange between distributed agents.
- Develop Web-based meta-data querying and integration tools. The suite of tools will enable end-users (a) to ascertain and compare meta-data formats from multiple

database sources and (b) to perform some translations between differing meta-data that will facilitate queries across heterogeneous data sources. These tools will be made available to end-users via the Web.

- Develop an ontology repository using the Object Management Group's Meta Object Facility (MOF). The work will be done in collaboration with DSTC and will use DTSC's MOF-based tools.
- Develop a prototype agent-based workflow management development system that employs the NZDIS framework and makes use of DSTC's FlowMake workflow experience.
- Develop enhanced agent-based security procedures for remote data access. This work will be done with an eye towards future DIS applications in electronic commerce.
- Develop an agent-based DIS framework over ATM networking. This will involve an adaptation of the lower-level agent platform to ATM communication and will be performed in collaboration with the work of Objective 2.
- Establish installations of all the tools as a pilot testbed for evaluation in a heterogeneous distributed network. This will allow users to pose realistic queries across data hosted by Landcare Research, GNS, and others. The pilot will be developed in consultation with a pan-industry focus group who will provide feedback on the relevance to New Zealand industry of the issues being addressed by the research group.
- Conduct annual national workshops, in collaboration with DSTC, at major urban centres that (a) present DISABC programme developments (b) present the latest national and international developments concerning DIS technology, and (c) gather direct end-user feedback concerning DIS issues.
- Provide a national Web portal to DIS technology, with both national and international links of interest. This will include online tutorials and other information associated with DISABC programme developments.

8. CONCLUSIONS

Accessing distributed information in an integrated fashion is difficult, because information stores are often distributed across various remote sites, stored on different platforms and in different formats, and organised according to differing organisational schemas and semantic models. The overall goal of the Distributed Information Systems (DIS) programme has been to develop advanced technology that addresses these difficulties.

A framework of requirements, from high level, overarching requirements down to the needs for specific tools was developed that forms a foundation for research and commercialisation. Key requirements include:

- to be able to integrate systems with differing meta-data;
- to be able to integrate systems with differing storage formats and media;
- to be able to incorporate legacy systems and unstructured datasets;
- customisable, to suit different application domains.

From the public and private sector consultations the research found that there is a genuine strong need for distributed technologies. Factors that require further work are concerned with:

- the need for meta-data standards,
- more flexible environments for integrating heterogeneous data sets,
- greater attention to security and access control.

The DIS project has made significant progress towards new DIS technologies and the results have been communicated to public and private sectors as well as being made available as public domain software.

The next few years will see a continued major expansion in the use of distributed information systems, and the technologies developed in this project will continue to play a part.

BIBLIOGRAPHY

DIS Papers Produced by the Project Team

Purvis, M., Cranefield, S., and Nowostawski, M., “A Distributed Architecture for Environmental Information Systems” in *Environmental Software Systems – Environmental Information and Decision Support,* Kluwer Academic, Dordrecht, The Netherlands (2000) 49-56.

Bush, G., Purvis, M., and Cranefield, S., “Experiences in the Development of an Agent Architecture”, in *Design and Applications of Intelligent Agents*, C. Zhang & V-W Soo (eds.), Springer-Verlag Lecture Notes in Artificial Intelligence, vol. 1881, ISBN: 3540679111, Berlin, Germany (2000) 76-87.

Huang, Z., Cranefield, S., Purvis, P. and McDonald, J. R, “A Proposal for a New Generation Hypertext Transfer Protocol on ATM Networks”, Chapter 34, *World Wide Web: Technologies and Applications for the New Millennium*, Computer Science Research, Education, and Applications Press, Athens, GA, U.S.A. (2000) 255-259.

Cranefield, S. and Purvis, M., “An agent-based architecture for software tool coordination”, *Intelligent Agent Systems: Theoretical and Practical Issues*, L. Cavedon, A. Rao, W. Wobcke (eds.), Springer-Verlag Lecture Notes in Artificial Intelligence, vol. 1209, Berlin, Germany (1997) 44-58.

Purvis, M., Zhou, Q., Cranefield, S., Ward, R., Raykov, R., and Jessberger, D., “Spatial Information Modelling and Analysis in a Distributed Environment” accepted for publication in *Environmental Modelling and Software* (2001).

Diaz, A., Cranefield, S., and Purvis, M., “Planning and Matchmaking in a Multi-Agent System for Software Integration”, *Proceedings of the 11th International Conference on Mathematical and Computer Modelling*, International Association for Mathematical and Computer Modelling (1997)

Nowostawski, M., Purvis, M. and Cranefield, S. Agent-Oriented Modelling for Complex Systems, in Applied Complexity—from neural nets to managed landscapes, S. Halloy and T. Williams (eds.), Christchurch, New Zealand Institute for Crop and Food Research (2000) 205–220

Huang, Z., Sun, C., Purvis, M., Cranefield, S., "Weak sequential consistency models for distributed shared memory", to appear in *Proceedings of the 10th International Conference on Computing and Information (ICCI '2000),* Kuwait (2000).

Cranefield, S., Purvis, M., and Nowostawski, M., “Is it an Ontology or an Abstract Syntax? Modelling Objects, Knowledge and Agent Messages”, *Proceedings of the Workshop on*

Applications of Ontologies and Problem-Solving Methods at the 14th European Conference on Artificial Intelligence (ECAI'00), Berlin, Germany (2000) 16.1–16.4.

Huang, Z., Sun, C., Purvis, M. and Cranefield, S. "View-based Consistency for Distributed Shared Memory", in *Proceedings of the Joint Meeting of the 4th World Multiconference on Systemics, Cybernetics and Informatics (SCI'2000) and the 6th International Conference on Information Systems Analysis and Synthesis (ISAS'2000), Volume VIII*, Lee, J. K., Juric, M., Bruzzone, A., Klovshy, D. and Fujita, M. (eds.) International Institute of Informatics and Cybernetics (2000) 1–6

Bush, G., Nowostawski, M., Cranefield, S. and Purvis, P. Platforms for Agent-Oriented Software Engineering, in Proceedings of the 7th Asia Pacific Software Engineering Conference (APSEC 2000), Dong J. S., He, J. and Purvis, M. (eds.), Los Alamitos, CA, USA, IEEE Computer Society Press (2000) 480–488

Cranefield, S. and Purvis, M., "Integrating Environmental Information: Incorporating Metadata in a Distributed Information Systems Architecture", to appear in *Proceedings of the Workshop Symposium on Integration in Environmental Information Systems (ISESS 2000)*, Environmental Informatics Institute, Gaiberg, Germany (2000).

Nowostawski, M., Purvis, M., and Cranefield, S., "An architectural approach to messaging in multi-agent systems", *Proceedings of the Autonomous Agents 2000 Workshop on Infrastructure for Scalable Multi-agent Systems (W13)*, Barcelona, Spain (2000) 105-110.

Cranefield, S. and Purvis, M., "Extending Agent Messaging to Enable OO Information Exchange", *Cybernetics and Systems 2000*, (Proceedings of the 15th European Meeting on Cybernetics and Systems Research), R.Trappl (*ed.*), Vienna, Austrian Society for Cybernetic Studies, Vienna, Austria (2000) 573-578.

Purvis, M., Cranefield. S., Bush, G., Carter, D., McKinlay, B., Nowostawski, M., and Ward, R., "The NZDIS Project: an Agent-based Distributed Information Systems Architecture", *Proceedings of the Hawai'i International Conference on System Sciences (HICSS-33)*, R. H. Sprague, Jr. (ed.), (CD ROM) IEEE Computer Society Press, Los Alamitos, CA (2000).

Huang, Z., Cranefield, S., Chee, V. K. M., Purvis, M., "A Java Networking API for ATM Networks", in *Proceedings of the 31th International Conference and Exhibition on Technology of Object-Oriented Languages and Systems (TOOLS Asia'99)*, Nanjing, China, (1999) 306-315.

Cranefield, S. and Purvis, M., "UML as an ontology modelling language" *Proceedings of the Workshop on Intelligent Information Integration, 16th International Joint Conference on Artificial Intelligence (IJCAI-99)*, (1999) 46-53.

Cranefield, S.J.S., Moreale, E., McKinlay, B. and Purvis, M. K., "Automating the Interoperation of Information Processing Tools", Proceedings of the 32nd Hawaii International Conference on System Sciences (HICSS-32). Maui, Hawaii, IEEE (CD-ROM) (1999) 10 pages.

Cranefield, S.J.S., McKinlay, B., Moreale, E. and Purvis, M.K. "Automating Information Processing Tasks: An Agent-based Architecture", *TZI Report 9/98: Proceedings of the Workshop on Intelligent Agents in Information and Process Management*, *22.* Deutsche Jahrestagung für Künstliche Intelligenz (22nd German Conference on Artificial Intelligence, KI-98). Bremen, Center for Computing Technologies (TZI), University of Bremen (1998).

Purvis, M., Cranefield, S., and Ward, R., "Distributed Software Systems: From Objects to Agents", *Software Engineering: Education & Practice*, Proceedings of the 1998 International Conference, IEEE Computer Society Press, Los Alamitos, CA (1998) 158-165.

DIS Reports Produced by the Project Team

Yu, Byung Hyun, *Design and Implementation of a Native Java Interface for ATM Network Applications*, M. Sc. Thesis, University of Otago (2000).

Nowostawski, M., Bush, G., Purvis, M. K., and Cranefield, S., "Platforms for Agent-oriented Software" *Information Science Discussion Paper Series*, Number 2000/13, ISSN1172-6024 (2000).

Cranefield, S., Purvis, M., and Nowostawski, M., "Is it an Ontology or an Abstract Syntax? Modelling Objects, Knowledge and Agent Messages", *Information Science Discussion Paper Series*, Number 2000/08, ISSN1172-6024 (2000).

Cranefield, S. and Purvis, M., "Extending Agent Messaging to Enable OO Information Exchange", *Information Science Discussion Paper Series*, Number 2000/07, ISSN1172-6024 (2000).

Cranefield, S. and Purvis, M., "Integrating Environmental Information: Incorporating Metadata in a Distributed Information Systems Architecture", *Information Science Discussion Paper Series*, Number 2000/02, ISSN1172-6024 (2000).

Purvis, M., Cranefield, S., Nowostawski, M., Bush, G., Carter, D., McKinlay, B., Ward, R., "The NZDIS Project: an Agent-based Distributed Information Systems Architecture", *Information Science Discussion Paper Series*, Number 99/17, ISSN1172-6024 (1999).

Purvis, M., Cranefield. S., and Nowostawski, M., "A Distributed Architecture for Environmental Information Systems", *Information Science Discussion Paper Series*, Number 99/06, ISSN 1172-6024 (1999).

Cranefield, S. and Purvis, M., "UML as an Ontology Modelling Language", *Information Science Discussion Paper Series*, Number 99/01, ISSN 1172-6024 (1999).

Cranefield, S., McKinlay, B., Moreale, E., Purvis, M., "Automating Information Processing Tasks: an Agent-based Architecture", *Information Science Discussion Paper Series*, Number 98/7, ISSN 1172-6024 (1998).

Useful Papers and Books on DIS Technologies

Brendon Cahoon, Kathryn S. McKinley and Zhihong Lu. "Evaluating the performance of distributed architectures for information retrieval using a variety of workloads". ACM Trans. Inf. Syst. 18, 1 (Jan. 2000), Pages 1 – 43.

S.-H Gary Chan and Fouad Tobagi. "Distributed servers architecture for networked video services". IEEE/ACM Trans. Networking 9, 2 (Apr. 2001), Pages 125 – 136.

Conrad T. K. Chang and Bruce R. Schatz. "Performance and implications of semantic indexing in a distributed environment". Proceedings of the eighth international conference on Information knowledge management, 1999, Pages 391 – 398.

Waiman Cheung and Cheng Hsu. "The model-assisted global query system for multiple databases in distributed enterprises". ACM Trans. Inf. Syst. 14, 4 (Oct. 1996), Pages 421 – 470.

Miyi Chung, Ruth Wilson, Kevin Shaw and Maria A. Cobb. "Distributing Mapping Objects with the Geospatial Information Database". Proceedings of the International Symposium on Distributed Objects and Applications, published by the IEEE, 1998.

Paul Dourish. "Using metalevel techniques in a flexible toolkit for CSCW applications". ACM Trans. Comput.-Hum. Interact. 5, 2 (Jun. 1998), Pages 109 – 155.

Jeri Edwards. "3-Tier Client/Server at Work". Pub. John Wiley, 1997.

Jacob Harris and Vivek Sarkar. "Lightweight object-oriented shared variables for distributed applications on the Internet". Proceedings of the ACM conference on Object-oriented programming, systems, languages, and applications, 1998, Pages 296 – 309.

Nigel Hinds, Chinya V. Ravishankar. "Managing Metadata for Distributed Information Servers". Proceedings of the 31st Hawaii International Conference on System Sciences (HICSS'98), published by the IEEE, 1998.

Nicholas Kassem. "Designing Enterprise Applications with the Java 2 Platform, Enterprise Edition". Pub. Addison-Wesley, 2000.

Oscar Luiz Monteiro de Farias and Luiz Má C. P. M. de Farias. "Distributed information system on an Internet/intranet environment (DISI2E)".Proceedings of the eighth ACM symposium on Advances in geographic information systems, 2000, Pages 197 – 198.

Julio C. Navas and Michael Wynblatt. "The network is the database: data management for highly distributed systems". To be published in: ACM SIGMOD international conference on Management of Data on Management of data, 2001.

Okada, R.; Eun-Seok Lee; Shiratori, N. "Agent-based approach for information gathering on highly distributed and heterogeneous environment". Proceedings of the 1996 International Conference on Parallel and Distributed Systems (ICPADS '96), published by the IEEE.

Robert Orfali and Dan Harkey. "Client/Server Programming with JAVA and CORBA". Pub. John Wiley, 1998.

Robert Orfali, Dan Harkey and Jeri Edwards. "Instant CORBA". Pub. John Wiley, 1997.

Robert Orfali, Dan Harkey and Jeri Edwards. "The Essential Distributed Objects Survival Guide, 2^{nd} Edition". Pub. John Wiley, 1996.

Owen de Kretser and Alistair Moffat. "Methodologies for Distributed Information Retrieval". Proceedings of the The 18th International Conference on Distributed Computing Systems, published by the IEEE, 1998.

Andreas Paepcke, Chen-Chuan K. Chang, Terry Winograd and Héctor García-Molina. "Interoperability for digital libraries worldwide". Commun. ACM 41, 4 (Apr. 1998), Pages 33 – 42.

Ahmed Saleh and George R. R. Justo. "A configuration-oriented framework for distributed multimedia applications". Proceedings of the 2000 ACM symposium on Applied computing 2000 (volume 1), 2000, Pages 278 -280.

Wolfgang Theilmann and Kurt Rothermel. "Disseminating Mobile Agents for Distributed Information Filtering". Proceedings of the First International Symposium on Agent Systems and Applications Third International Symposium on Mobile Agents, published by the IEEE, 1998.

Edgar Weippl. "Building secure knowledge bases: combining Java agents and Dbagents". To be published in: Proceedings of the fifth international conference on Autonomous agents, 2001.

Jinxi Xu and W. Bruce Croft. "Cluster-based language models for distributed retrieval". Proceedings of the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, 1999, Pages 254 – 261.

Zhili Zhang and William Perrizo. "Distributed query processing using active networks". Proceedings of the 2000 ACM symposium on Applied computing 2000 (volume 1), 2000, Pages 374 – 380.