National Research FLAGSHIPS
Water for a Healthy Country

CSIRO

Web-based Visualisation of Water Information: an overview

R.G. O’Hagan, B. Robinson, G. Swan

19 March 2008

Bureau of Meteorology
Dr. D. Kinny

Water for a Healthy Country Flagship Report series ISSN: 1835-095X

Australia is founding its future on science and innovation. Its national science agency, CSIRO, is a powerhouse of ideas, technologies and skills.

CSIRO initiated the National Research Flagships to address Australia's major research challenges and opportunities. They apply large scale, long term, multidisciplinary science and aim for widespread adoption of solutions. The Flagship Collaboration Fund supports the best and brightest researchers to address these complex challenges through partnerships between CSIRO, universities, research agencies and industry.

The Water for a Healthy Country Flagship aims to achieve a tenfold increase in the economic, social and environmental benefits from water by 2025. The work contained in this report is collaboration between CSIRO and the Bureau of Meteorology.

For more information about Water for a Healthy Country Flagship or the National Research Flagship Initiative visit www.csiro.au/org/HealthyCountry.html

Citation: O'Hagan, RG, Robinson, B, Swan, G and Finny, D, 2008. Web-based visualisation of water information: an overview. CSIRO: Water for a Healthy Country National Research Flagship, 57 pp.

Enquiries should be addressed to:
Garry Swan
CSIRO Land and Water

Copyright and Disclaimer

Important Disclaimer

Contents

List of Figures

List of Tables

EXECUTIVE SUMMARY

The **WIRADA Bridging projects** are a set of preliminary studies contracted to CSIRO to contribute to the direction and development of the Bureau's Water Information System specifically focussing on the creation and delivery of data. The overall objective of these studies is to help the newly formed Water Division to prepare the development of the future systems and methodologies to deliver a national water account and water assessment products. The Data Delivery project consists of several activities investigating various aspects of data delivery. These activities include issues related to delivery of observational and gridded data, standards for water information exchange, and presentation of water information data. The focus of this activity is the presentation of water information data.

The Australian Water Resources Information System (AWRIS) will require the development of end user web based applications for the presentation and reporting of water data. The reporting tools component of the Data Delivery Project will need to undertake a stakeholder/end user requirements analysis (from software engineering perspective) of the visualisation and information presentation needs for AWRIS. The outcome of this activity is to provide an overview of various technologies and processes that AWRIS can use to support user requirements in the area of reporting applications.

The three key areas where visualisation and information presentation are fundamental to AWRIS are Information Discovery, Information Portrayal and Information Analysis. Information Discovery refers to facilitating users in identifying information of interest. Information Portrayal is concerned with presenting information in ways that enhance understanding, and Information Analysis is about allowing users to investigate data, relationships and trends. These top-level use cases can each be broken down into more specific visualisation and reporting tasks. Information discovery allows users to identify information of interest from the mass of water information that will be available. Interfaces that allow effective discovery of pertinent information require an understanding of the types of questions users want answered. This understanding can then be used to develop query interfaces suited to user needs. Information presentation refers to the portrayal of data in ways that provide information and enhance understanding. The method of information presentation will vary depending on the type of data and the questions posed by the user. In the case of water information, presentation methods will primarily need to focus on display of spatio-temporal data. The final key use case, information analysis, refers to the provision of tools for investigating data – for example by graphing to identify relationships and trends or performing statistical analyses.

In this report, key components associated with the use cases identified above are explored. This includes query interfaces, spatial or map representation of data, 3D data visualisation, portrayal of temporal data, graphing and data analysis using data cubes. For each of these areas, an overview of the current state of the art is provided along with details of significant techniques or technologies involved. Details of particular vendor implementations are largely avoided as implementation choice can be considered independently from choice of methodology.

In evaluating all the component areas in this report, several general observations can be made. First, the design process for AWRIS web applications should include a detailed analysis of the requirements of potential user groups. Understanding why users will use the system, and what

they will use it for, will be the key in determining how the system should be developed. This understanding will underpin all components – the query interface, how map displays can be best used, what temporal information will need to be portrayed and how, and how best to utilise technologies such as data cubing to optimise data retrieval. Secondly, water resource information is inherently spatial in nature, with a strongly associated temporal component. Observational data will form a large proportion of the total data holdings, and so interfaces and visualisations need to be developed within the spatio-temporal context. Lastly, following on from the user requirements analysis suggesting in the first point, an in-depth investigation of presentation techniques should be carried out based on the key questions and data types. This would include evaluation and prototyping of various ways to portray the selected information sets in order to address specific queries. This targeted approach for specific water information scenarios would enable a more effective design of the system in the long term. It should also be recognised that the needs of various users may differ, resulting in different levels of requirements from simple services to find appropriate data sets to complex access to data analysis. The deployed system will need to be stratified to support the different communities of users.

1. INTRODUCTION

1.1 Objectives of the WIRADA "bridging projects"

The **WIRADA Bridging projects** are a set of preliminary studies contracted to CSIRO to contribute to the direction and development of the Bureau's Water Information System specifically focussing on the creation and delivery of data. The overall objective of these studies is to help the newly formed Water Division to prepare the development of the future systems and methodologies to deliver a national water account and water assessment products. Such systems will need to create, process, and serve a variety of different data types including:

- GeoFabric Data (Geospatial Features).
- Observation Data such as flow and height measurements;
- Gridded data such as interpolated rainfall and ET surfaces;

This state of the art review focuses mainly on features and observations to feed two separate projects:

- The Australian Hydrologic Geospatial Fabric (AHGF) project
- The Data Delivery project: this review is a sub-activity of the WaterML work package of this project.

1.1.1 The Australian Hydrologic Geospatial Fabric (AHGF) project

The purpose of the AHGF project is to define the content and function of the first release of the Australian Hydrological Geospatial Fabric to meet the needs of the Australian Water Resources Information System (AWRIS) (AWRIS 2005) and the wider community. The activity will provide the specification for each data set along with recommendations for how it might be developed and its ongoing governance.

An important element of a national water resources information system is knowledge of the features within that system: the catchments, streams, aquifers, floodplains, storages, and wetlands that make up the hydrological system. To develop and deliver water resources information, location of these features is vital as is information about the interactions between features. At present there is no nationally consistent hydrological data set for Australia at a scale larger than 1:250,000. That is, there is no consistent stream network, the national DEM is of limited value for hydrological forecasting and accounting and data sets maintained by State agencies cannot simply be joined to produce a national picture. Furthermore, no single agency has, to date, taken responsibility for such a data set.

The development of an **Australian Hydrological Geospatial Fabric** will allow all those working in water management and research to refer to a single, consistent, national geospatial framework for hydro features. The AHGF project activities are centred on:

- AHGF Specification: Identify the requirements of an Australian Hydrological Geospatial Fabric: describing the key data sets and their nature and relationships; and,
- Arc Hydro Pilot: Test this design through an implementation of the Arc Hydro toolset to deliver a simple water balance for a hydrologically complex region in Australia;

1.1.2 The Data Delivery project

This Data Delivery project includes a set of activities that will contribute to the direction and development of the Bureau's Water Information System specifically focussing on the creation and delivery of data. The outputs from these activities will be recommendations in the form of technical reports to the Bureau. In the context of meeting the short term needs for AWRIS, the project activities will:

- Assess approaches to the web service delivery of observational and gridded data.
- Assess current standards suitable for use as a water information exchange format that can be adopted by agencies who will become data providers to the Bureau.
- Investigate methodologies for the creation of a scalable reusable reporting 'dashboard' or end user applications and tools.

1.2 Data Delivery Project Work Breakdown Structure

The work packages of the Data Delivery project are:

WP 1.1. Assessment of observational data services (AWDIP and HIS): the Australian Water Data infrastructure project (AWDIP) as an example of an observational point of truth data service, has progressed over the last 3 years to a point where implementation of services have been deployed (as proof of concept) to most states in Australia. Work package 1.1 will provide advice and recommendations to the Bureau on the opportunity to reuse and extend the work done in AWDIP vs. alternative implementations such as CUAHSI's HIS Server.

WP 1.2. Evaluation of Delft FEWs for data delivery: Delft FEWs, is an application suite developed by Delft Hydraulics in the Netherlands as an operational Flood Early Warning system (FEWs). Work package 1.2 will investigate and assess the data management strategies deployed by Delft for use in an AWRIS/AWDIP context and make recommendations to the Bureau on which are the salient aspects of this system.

WP 1.3. Assessment of gridded data services: Gridded data will be required by a number of modelling applications by the Bureau. Work package 1.3 will assess implementations of gridded data services by working on a range of concrete examples, such as the AWAP project.

WP 2. Towards defining a water data exchange format (WaterML): Defining a standard water information exchange format is one of the fundamental tasks for the Bureau in the establishment of a water information system. This Work Package 2 will review the state of the art on water standards to guide the development of a conceptual model of water resources information, corresponding markup language (ML) and future evaluation activities. A secondary objective of WP2 is to foster international collaboration on this topic.

WP 3. Water Reporting Tools: AWRIS will require the development of a data dashboard (end user web based applications) for the presentation and reporting of water data. Work package 3 will undertake a stakeholder/end user requirements analysis (from software engineering perspective) for AWRIS data dashboard and data explorer product suite. The outcome of this activity will be a technical report that details the suitability of available products and technologies that AWRIS can use to develop data reporting applications.

1.3 Change to Original Objective

The original objective of work package 3 was:

"To identify the ultimate makeup of the AWRIS Data Dashboard, data explorer and data download service to a point where detailed specifications can be prepared for commercial build or possible delivery by existing commercial toolsets" (Stewart, 2008).

At the commencement of this project, an evaluation of AWRIS user requirements for a data dashboard was attempted. It rapidly became clear that existing requirements were very high level, lacking the detail required to understand what the dashboard would provide to users. Attempts to elicit more detailed requirements from potential users and designers of the system were largely unsuccessful – primarily because users and designers from the Bureau were either unavailable for consultation, or were unable to articulate how the proposed system would be used.

Due to the lack of progress in determining requirements for the AWRIS data dashboard, and the limited time available for completion of the project, the objective of this activity was modified to being a more general evaluation of tools and technologies suitable for the presentation (and in particular, the visualisation) of water resource information.

1.4 Scope of work package 3 and of this state of the art review

This work package aims to explore the expected end user requirements for AWRIS, and in particular, to look at some of the technologies and techniques involved in delivering water information over the web.

It should be understood that the scope had to be inferred from document analysis as the AWRIS development at the stage of writing this report was still being put together. Analysis of AWRIS documentation has provided a high level scope from which usage was inferred. This document provides an insight to the types of issues the AWRIS team will need to be prepared for when designing systems.

The purpose of AWRIS is the provision of tools that allow ongoing access to distributed sources of water resource information. This will contribute to ensuring consistency in the understanding and reporting of water information, and facilitate the delivery of water assessments, accounts and other reports. AWRIS is expected to have a broad user community encompassing domains such as policy and government, research, environmental planning and management, operational agencies and the general public. Each element of this varied group of users will have their own reasons for using the system, and their own requirements for access to, and use of, the information available. AWRIS needs to be designed such that it accommodates as many of these requirements as possible.

This report identifies the areas in which visualisation will contribute to AWRIS and the requirements for those areas. It also provides an evaluation of the current 'state of the art' in presentation of information over the internet, and details issues involved. The report is intended to provide an overview of technologies and processes rather than a detailed evaluation of specific tools or vendor implementations. The report is structured as follows: Section 2 proposes key user requirements for visualisation of water information, and discusses the high-level use cases that can be determined from these requirements. Section 3 then examines some of the major technology components that may be implemented to support the proposed user requirements and also describes processes that may be used to satisfy the derived use cases. This section examines web architectures, query interfaces, portrayal techniques such as mapping, 3D visualisation and visualisation of spatio-temporal information, and analysis tools such as graphing and data cubes. The conclusion of the report is in Section 4 where recommendations are summarised.

2. VISUALISATION USER REQUIREMENTS FOR AWRIS

Current AWRIS documentation (NWC, 2006) includes a brief analysis of user requirements derived from interviews of potential users. This analysis provides a useful starting point, but is limited in that the AWRIS system and the data sets it will support was in the very early stages of design at the time of the interviews. A full, more detailed analysis is needed in order to identify the key scenarios in which AWRIS will be used, and hence the requirements of particular user communities. As this document is based on the initial analysis, it provides an insight to the types of issues the AWRIS team will need to be prepared for when designing systems, but does not reflect the level of detail required to successfully meet end-user requirements.

2.1 Deriving User Requirements

The AWRIS vision states that "AWRIS will provide a comprehensive, credible, open view of Australia's water resources data and information. It will allow users to review and investigate this data, understand the data in context, and bigger picture details about the data" (NWC, 2006). This vision can be decomposed into a set of high level requirements for AWRIS:

- Provide a comprehensive view of Australia's water resources data and information;
- Provide a credible view of Australia's water resources data and information;
- Provide open access to Australia's water resources data and information;
- Allow review of data;
- Allow investigation of data; and
- Provide context for the data.

From a visualisation perspective, these requirements indicate that the system must allow users to not only view available data, but also to analyse and explore data within a broader context. This in turn provides insight into the role and key use cases of visualisation and reporting within AWRIS.

2.2 AWRIS Visualisation Use Cases

The key, high-level, use cases for visualisation within AWRIS can be generalised as being part of the following concepts:

1. Information discovery – allowing users to identify information of interest;
2. Information portrayal – presenting information in ways that enhance understanding; and

3. Information analysis – allowing users to investigate data, relationships and trends.

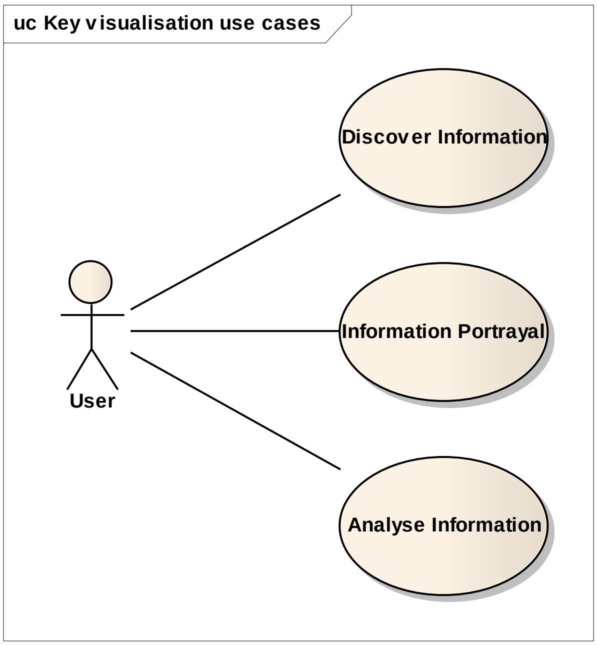

Figure 1 Key Visualisation Use Cases

Exploration and detailing of these three key use cases will elicit specific requirements for visualisation and reporting of water information.

2.3 Information Discovery

AWRIS will provide a comprehensive view of Australia's water resources information. This means that users will have access to vast amounts of data on, for example, the location, quantity, quality and use of water at national, regional or local scales. Allowing users to identify data of particular interest is critical to the efficacy of the system. The Information Discovery use case encompasses presentation of (large amounts of) information to facilitate discovery and selection of relevant data sets. Given that the majority of water information is inherently spatial in nature, spatial representation of information will play a large role in discovery and selection of available resource information.

In order for users to discover particular sets of information, they must be able to pose questions of the system. These questions may be specific, for example "What rainfall and river flow information is available for the Warrego catchment between 1920 and 2006?", or they may consist of a set of questions that drill down through available data until the desired data set is obtained:

- "I want rainfall and river flow information"
- "I want it for the Warrego catchment"

- “I want data between 1970 and today”

The types of questions that users ask greatly inform the way in which information is best presented to the user. While it is unlikely that a single user interface will be appropriate for every possible question, its design should be directed based on the set of questions most likely to be asked.

At a high level, these questions may be of the form:

- What data is available?
- What is the spatial extent or coverage of the data (latitude and longitude coverage)?
- What is the temporal coverage of the data (start date/time and finish date/time)?
- In what ways can the data be viewed?

Subsequent questions comparing one data set with others, or questions regarding the content of the data will be addressed in Section 2.5 on information analysis.

The questions described above can be used to derive a set of use cases for Information Discovery. In order for the user to understand what data is available, the system will need to provide a view of data sets – either by theme, spatial or temporal coverage. The user will then select information based on theme, location/region of interest or temporal coverage from the available information provided by the system, and choose in which way the information should be presented (determine the format of the information). These use cases are shown in Figure 2.

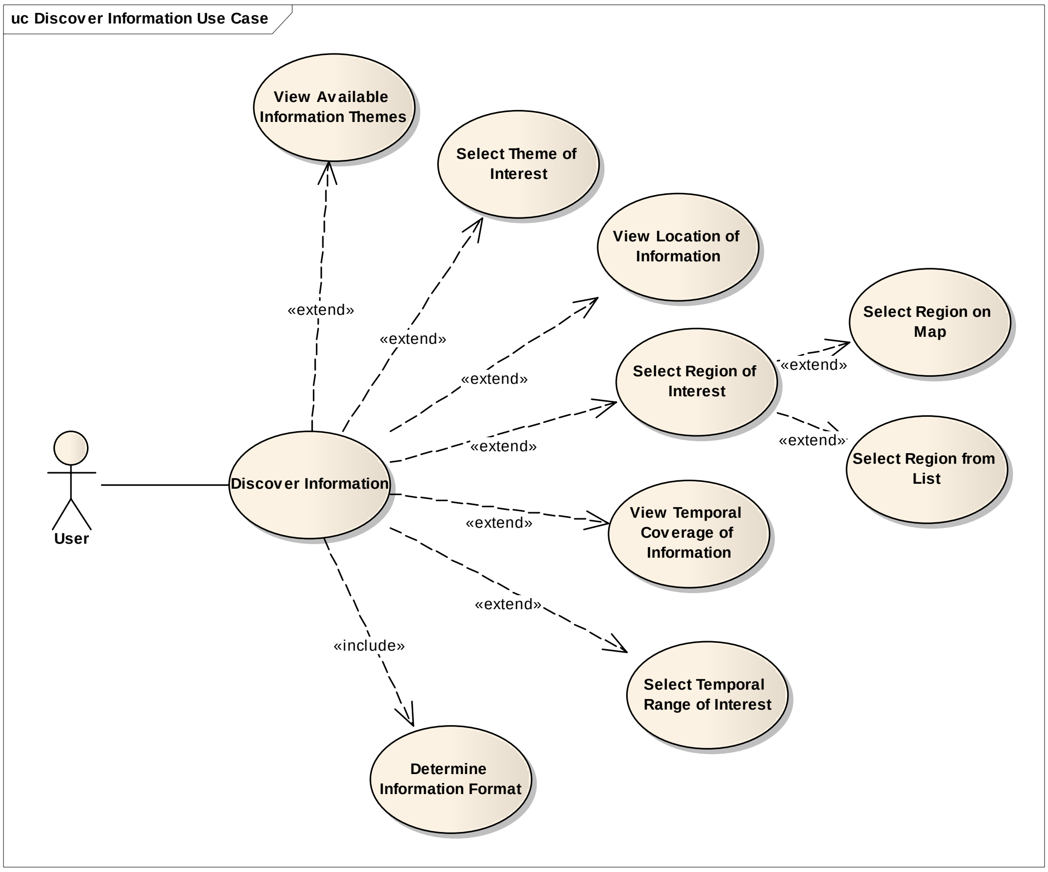

Figure 2 Information Discovery Use Case

In order to enable the user to ask the types of questions posed above, AWRIS will require a query interface that allows the user to:

- spatially select a region of interest;
- specify one or more variables or domains of interest, for example rainfall, stream flow, salinity, pH, entitlements;
- select a temporal range of interest; and
- select the way in which resulting data sets are presented.

AWRIS will also require the ability to apply questions to its data holdings, and return appropriate results. This will require, at the least, computational metadata for each data set, and possibly the incorporation of semantic reasoning. The way in which the system handles and responds to user-defined queries is beyond the scope of this document. The way in which results are presented to the user is within scope, and is discussed in the following section on Information Portrayal.

2.4 Information Portrayal

Presentation of water resources information is a fundamental part of AWRIS (see Section 2.1). Information visualisation will be needed to support a number of user activities:

- To show users what data is available and assist in data selection;
- To provide summaries of water information;
- To display detailed water information; and
- To provide reports.

To satisfy the above visualisation demands, AWRIS will need a variety of different forms of information presentation. The method of presentation chosen will depend on the type of information that will be portrayed, as well as the purpose or audience of the visualisation. Types of information representation can be expected to include more traditional maps, graphs and tables as well as more innovative methods such as animated surfaces showing time-series data or 3D models indicating the location, extent and topography of groundwater aquifers. For each of these representations, the content may be static (as in the case of pre-generated maps of rainfall from the past 24hrs), or dynamically generated in response to specific queries.

As an example, consider the types of information presentations contained in a State Water Report. The Victorian State Water Report for 2004/05 (Victorian Government 2006) contains information on rainfall, stream flow, surface water storage levels, ground water salinity and storages, entitlements, environmental flows, water availability and water use. This information is presented primarily in the form of maps, graphs and tables. The report presents the current value or status of these variables, and compares them with results from previous years or long term averages. Due to the static nature of the report document, the method of presentation and the information that can be included in the report is limited. These limitations are removed when considering online delivery as content can be generated dynamically and presentations produced on the fly. In an online version of a State Water Report, the user could compare current values with all available historic data rather than just the previous year, and animations could be used to display time series results.

2.5 Information Analysis

Although presentation of information to the user will enable enhanced understanding of water resources in Australia, display of static, pre-defined views of information will not fulfil all requirements. The ability of the user to ask different questions, to analyse information in new ways and request specific views of data is also important.

Information analysis involves interaction between the user and the system to dynamically generate new views of a data set, or results in the form of derived data. This may include the ability to perform computations on data, such as statistical methods, inputs to model data, and the subsequent presentation of results in the form of graphs, tables or other visualisations. It may also include the ability to directly interact with a visualisation, to slice a data set or generate a profile of a particular path or transect. Or, it may include the ability to compose a workflow of processes, filters or transformations that will be applied to a data set to generate new results.

While an information portal or data dashboard will generally provide the ability to view the results to common or expected questions, the flexibility and power of the system rests on its ability to respond to arbitrary, user-defined queries, and to provide a generic framework for reporting or analysis. These requirements will influence AWRIS design on the nature of data access provided to users, furthermore it implies that data access will need to be stratified to support the different user communities.

Figure 3 shows some of the use cases that make up information analysis.

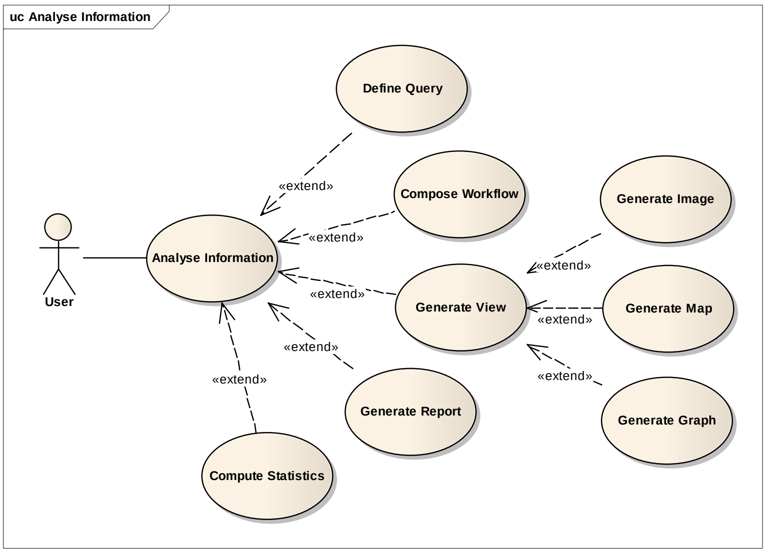

Figure 3 Information analysis use cases

3. VISUALISATION COMPONENT REQUIREMENTS

To meet the requirements of information discovery, portrayal and analysis that AWRIS is required to support, a number of components will be needed. An underlying framework that allows these components to form an integrated system for reporting and presentation of water information will also be required. The AWRIS User Requirements document (NWC, 2006) details some of these components. In this section, an overview is provided for components that support the use cases described in Section 2 above.

3.1 Web Application Architectures

Traditional web applications follow a ‘page submit and refresh’ model of interaction where user actions trigger HTTP requests which are sent back to a web server. The web server then processes the request (retrieving data, performing computation or interacting with legacy systems as necessary) and returns an HTML page to the client browser (Garrett 2005). This model is characterised by poor user experience and poor performance due to the high level of server access required, the delay in response as requests are handled synchronously, and the redundant refreshing of the entire page for each piece of new information (Henry 2007). This traditional model is shown in Figure 4.

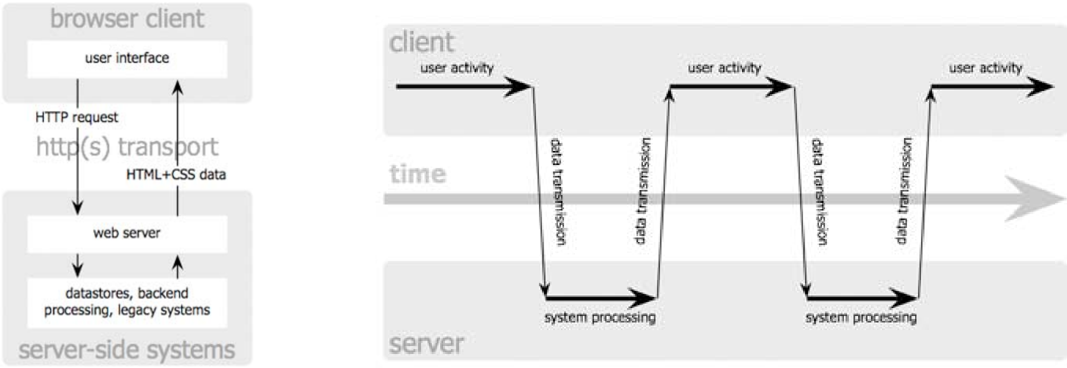

Figure 4 Classic web application model (source: Garrett, adaptivepath.com)

Rich Internet Applications (RIA) are a more recent paradigm that aim to provide the same level of interactivity and responsiveness as desktop applications (White). They differ from traditional web applications by introducing a client-side rendering engine that allows some processing to be carried out on the client. The client-side engine also enables asynchronous processing independent of communication with the back-end web server (Deb 2007). This in turn leads to richer, more powerful user interfaces, and a much higher level of responsiveness and interactivity. A typical RIA architecture is shown in

Figure 5. There are a number of RIA technologies and implementations including AJAX, Flash, and Java Applets which follow the architecture described here.

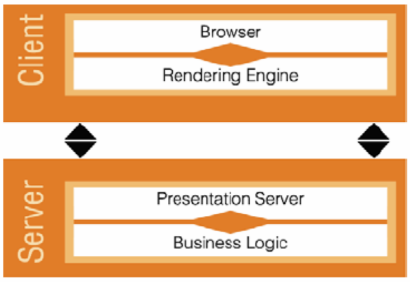

Figure 5 Typical Rich Internet Application architecture (source: Deb et. al. 2007)

The design of server-side systems is equally important as selection of a web application architecture. The web server should be developed in such a way as to ensure scalability, flexibility and extensibility. As indicated in Section 2.2, AWRIS will need to support a variety of types of information presentation. Transforming data from raw storage formats into particular presentation formats can be achieved through a set of transformation gateway services. This allows for new data services, source formats or presentation formats to be added to the system at any stage without affecting existing capabilities. An example of this can be seen in Figure 6 where data may exist in any of a number of different storage formats or technologies (e.g. files, databases).

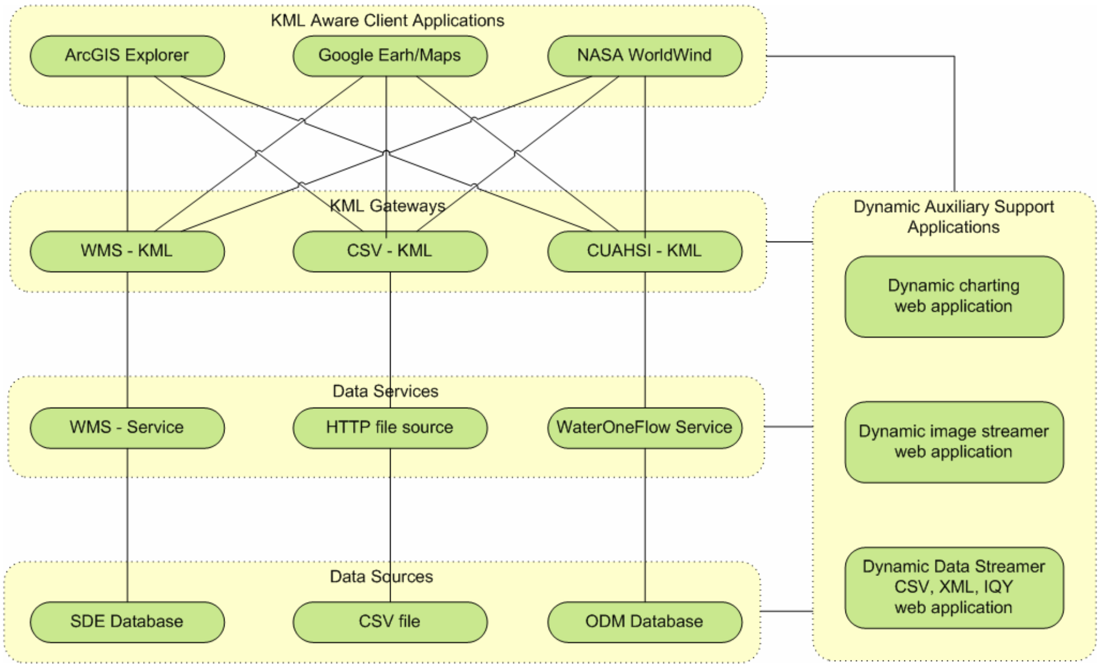

Figure 6 Web server design example showing layers and gateways

Each of these can also be served in a variety of ways using data services such as WMS, CUAHSI WaterOneFlow services or even simple HTTP file sources. Finally, the data can be presented using technologies such as Keyhole Markup Language (KML) via gateway services that transform the data formats from data services into a presentation format (KML) for consumption by the client application or browser. Auxiliary applications support the process by rendering the embedded content within the KML. Other gateways might transform data into RSS feeds or simple HTML pages (rather than KML) and would use the same auxiliary applications to render embedded graphs or images (for example). KML is described in more detail in Section 3.3.4.

3.2 Query interface

In order for the user to select particular elements or sets of information that are of interest, some method of forming questions against data holdings is required. The design of this query interface is critical to the usability of the system – users must be able to easily retrieve the information they require from the data available to them.

Query interfaces for spatial data typically consist of a combination of text- and map-based query components. This allows the user to spatially locate an area of interest on a map, while specifying non-spatial attributes or criteria through keywords or list selections. The way in which users interact with the map component and the text or selection components is determined by whether queries are expected to be location-based or theme-based. Location-based queries are of the form “What data is available for this region?”, for example “What are the current dam levels in the ACT?”, whereas theme-based queries are looking for information irrespective of location – “What stream flow data is available?” An optimal interface design for location-based queries may not be optimal for theme-based queries. As the amount of available data increases, this will become even more important. A clear understanding of user types and their requirements is necessary before designing the query interface.

Examples of query interfaces that combine map and text/selection components demonstrate the differences between location-focused and theme-focused interfaces.

Location-focused data repositories:

Geospatial Data Gateway – http://datagateway.nrcs.usda.gov/
This is a US Department of Agriculture site that provides a ‘one-stop source’ for natural resources or environmental data. The site provides a wizard style of interface that starts with selection of a region of interest, and then displays a tree-view of available data sets for that region. Once data sets are selected, the user can customise how the data will be delivered (e.g. format and data projection). In the final step, data sets are made available for download (FTP) or supply via CDROM or DVD.
The main disadvantage of location-focused interfaces such as this example, is that the data sets available are unknown until a region of interest is selected, precluding generic queries for data on a specific theme of interest.

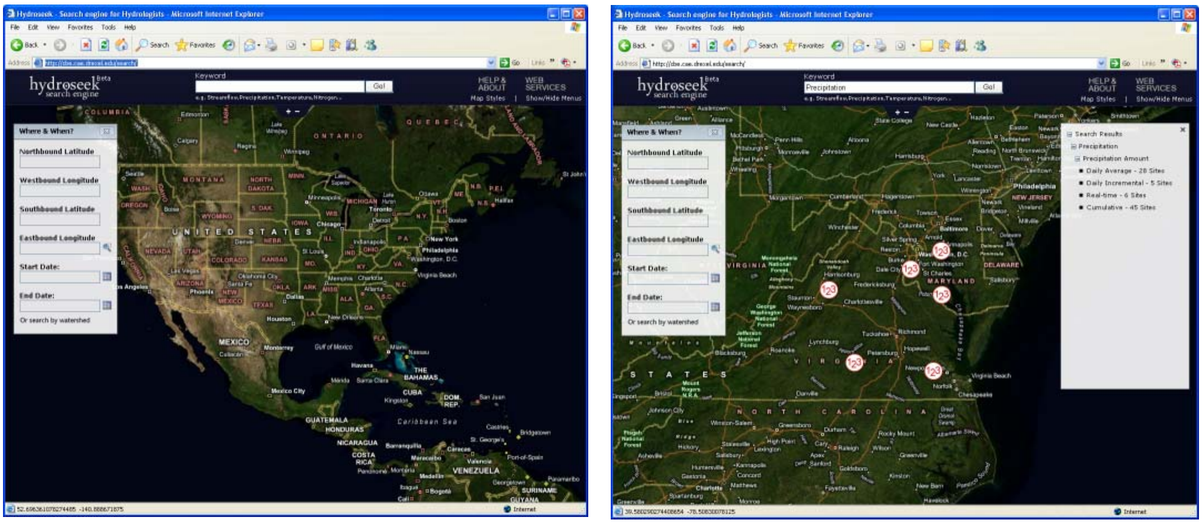

Figure 7 USDA's Geospatial Data Gateway – Step 1 (left): select region, Step 2 (right): select data set

Theme-focused data repositories:
Hydroseek – http://cbe.cae.drexel.edu/search/
Hydroseek is an ontology-aided search engine for US hydrology data providing a unified view over databases from a variety of agencies. The Hydroseek search engine is aimed at

hydrologists, and provides a combination map/text interface. The map component allows the user to interactively zoom in and out to an appropriate scale, and to select a region by bounding box or watershed. The text search can then be used to identify particular data sets by keyword. The location of specified data sets is then displayed on the map with summarised information when the map is zoomed out.

Figure 8 Hydroseek search engine website start page (left) & search for Precipitation – Daily Average (right)

One disadvantage of the Hydroseek query interface is that the user must use search terms that are within the vocabulary in order to find data. There is no list of these terms or the available data sets within the interface. This may be appropriate for users looking for specific information where a particular term is in common usage (e.g. streamflow), but is less helpful for users wishing to browse the data holdings or search for less standardised phrases (e.g. searching for ‘dam’ or ‘reservoir’ returned zero results). The interface does provide some assistance in the form of word choices once the first few letters have been entered.
It is likely that a combined theme- and location-focused interface will provide a solution that addresses the needs of different user groups rather than one or the other.

3.3 Mapping / Spatial Visualisation (2D)

Mapping / spatial visualisation (2D) refers to the portrayal of water data and information that has a geographic location in a way that displays the relationships between points in space. Map displays can be expected to form a large part of the visualisation requirements within AWRIS, both for discovery (spatial searches) and display of results.

The mapping or spatial visualisation component within AWRIS should be able to generate dynamic, interactive maps in response to user queries or requests. It should also be capable of displaying a range of different content over the background map image including image overlays, annotation, and various symbols or shapes.

Many systems now incorporate mapping components in order to present information spatially. It is recommended that AWRIS design incorporate Standards-based Components Off The Shelf (SCOTS) that support de jure standards, rather that those that support proprietary standards. The lead geo-spatial standards body is the Open Geospatial Consortium Inc. (OGC). The OGC is a non-profit, international voluntary consensus standards organisation that develops and promotes standards for geospatial and location-based services. The OGC has several standards and discussion papers relating to web map delivery for which overviews are provided below.

3.3.1 Web Map Service (WMS)

The OpenGIS Web Map Service (WMS) Implementation Specification provides operations “in support of the creation and display of registered and superimposed map-like views of information that come simultaneously from multiple remote and heterogeneous sources” (de la Beaujardiere 2004). The current version of this standard (version 1.3.0, 2006-03-15) is available at http://www.opengeospatial.org/standards/wms.

The WMS Specification describes three HTTP requests. The first is ***GetCapabilities*** which returns an XML document that describes the map layers available and the server's capabilities (i.e. the image formats, projections, and geographic bounds of the server.) An example GetCapabilities request to Geoscience Australia’s WMS looks like:

http://www.ga.gov.au/wms/getmap?dataset=national&request=getCapabilities&version=1.1.1&service=WMS

The second request is ***GetMap*** which returns a raster map image. The request arguments, such as the layer id and image format should match what is listed as available in the GetCapabilities return document. A sample GetMap request has the form:

http://www.ga.gov.au/wms/getmap?dataset=national&version=1.1.1&service=WMS&request=GetMap&layers=hydgeo_lm&format=image/png&WIDTH=512&HEIGHT=512&BBOX=-2100000,-5000000,1900000,-1000000&SRS=EPSG:3112

The third request, ***GetFeatureInfo***, is optional and is designed to provide WMS clients with more information about features in the map images that were returned by earlier GetMap requests. The response should contain data relating to the features nearest to an image coordinate specified in the GetFeatureInfo request. The structure of the data returned is not defined in the specification and is left up to the WMS server implementation.

There are a number of WMS Clients available, as both free and commercial products. Some of the more well known clients are listed below.

- Google Earth - http://earth.google.com/ (free)
- NASA World Wind - http://worldwind.arc.nasa.gov/ (open source, .net and java versions)
- OpenLayers - http://openlayers.org/ (open source)
- ESRI ArcGIS and ArcExplorer Web - http://www.esri.com

There are also multiple WMS Servers available including the following:

- GeoServer - http://geoserver.org (open source)
- Mapserver - http://mapserver.gis.umn.edu/ (open source)

- Oracle MapViewer - http://www.oracle.com/technology/products/mapviewer/index.html
- ESRI ArcIMS - http://www.esri.com

The OGC registered products pages at: http://www.opengeospatial.org/resource/products contains a more exhaustive list of products that have registered as supporting one or more OGC specifications including WMS.

3.3.2 Proposed Animation Service Extension

The proposed Animation Service Extension discussion paper (version 0.9, date: 2005-07-24) (LaMar 2005) is a document explaining how the Web Map Server specification (version 1.1.1 at time of writing) could be extended to allow animations that move in space over time (for example the tracking of a cyclone).

The WMS specification does not provide a method of serving animations as its primary focus is mapping with still imagery; however, it does provide a key element necessary for animations: the ability to associate a time with an image. To retrieve an animation from a standard WMS server, a client would submit a series of requests, each with a different timestamp constraint and then display the resultant images in order to show the animation.

The Animation Service Extension proposes adding two new request types, ***GetFrames*** and ***GetAnimation*** to the WMS. The GetFrames operation would allow requests for several animation frames at once, with the result being a multipart MIME document containing a descriptive XML document and a raster image for each frame requested. The GetAnimation operation would also return a multipart MIME document. In this case, instead of containing individual raster images, the animation sequence would be compressed using MSMPEG4 encoding into a single movie file (eg AVI or MPG). Both operations are designed to reduce the network communications overhead associated with making multiple GetMap requests. As an example, the following movie could be retrieved from a WMS server with a single GetAnimation request: *http://svs.gsfc.nasa.gov/vis/a000000/a003300/a003352/a003352_H264_512x512.mp4*

At present, implementations of this proposed animation extension are not available. It should also be noted that given the proposal was written in 2005, support for this extension appears to be limited.

3.3.3 Styled Layer Descriptor (SLD) Profile of the Web Map Service

The Styled Layer Descriptor (SLD) profile (Lupp 2007) explains how the Web Map Service specification can be extended to allow user-defined symbolisation of feature and coverage data. An integrated SLD-WMS is a basic WMS that allows clients to apply Symbology Encoding to all or a subset of its layers. In addition to the WMS operations of GetCapabilities, GetMap and GetFeatureInfo, an integrated SLD-WMS has to support the DescribeLayer and DescribeFeatureType operations.

3.3.4 KML 2.2 – An OGC Best Practice

Google submitted KML (formerly Keyhole Markup Language) as a Best Practice document to the Open Geospatial Consortium Inc. with the goal of adoption as an OGC Implementation Standard. The KML 2.2 Best Practice document (available from http://www.opengeospatial.org/standards/bp) specifies KML, an XML grammar used to visualise geographic data in an earth browser such as a 3D virtual globe and 2D web browser or mobile mapping applications.(Wilson 2007). KML is complementary to the WFS and WMS standards and utilises certain geometry elements derived from GML (Geography Markup Language). The key features of KML include the ability to:

- Annotate the Earth;
- Specify icons and labels to identify locations on the surface of the planet;
- Create different camera positions to define unique views for KML features;
- Define image overlays to attach to the ground or screen;
- Define styles to specify KML feature appearance;
- Write HTML descriptions of KML features, including hyperlinks and embedded images;
- Organize KML features into hierarchies;
- Locate and update retrieved KML documents from local or remote network locations; and
- Define the location and orientation of textured 3D objects.

There are a number of clients that support KML, the most well-known being GoogleEarth and GoogleMaps. Others include NASA's WorldWind and Microsoft's VirtualEarth, and ESRI's ArcGIS Explorer – all virtual world or 'geobrowser' applications. While geobrowser applications require downloading and installation on a user's computer, and hence are not really 'web applications', their use is becoming more and more widespread among the general user community, and so should be considered.

Generation of KML currently normally requires development of a gateway or transformation service, although applications such as ArcGIS Server are able to export data in KML directly. An example gateway has been developed by CSIRO Land & Water to deliver hydrological data from CUAHSI WaterOneFlow services in KML. This provides one presentation layer for data from WaterOneFlow services. The KML can then be delivered to a web browser via GoogleMaps or a geobrowser such as GoogleEarth. This example can be seen at http://wron.net.au/WebApps/HIStoKML/ODM2KMZ.aspx?URL=http://wron.net.au/CsiroHIS/cuahsi_1_0.asmx&Grouping=County&Show=All or by providing a WaterOneFlow service URL such as http://wron.net.au/CsiroHIS/cuahsi_1_0.asmx to the CSIRO gateway service at http://wron.net.au/WebApps/HIStoKML/ODM2KMZGateway.aspx. The resulting KML file can then be loaded into GoogleEarth.

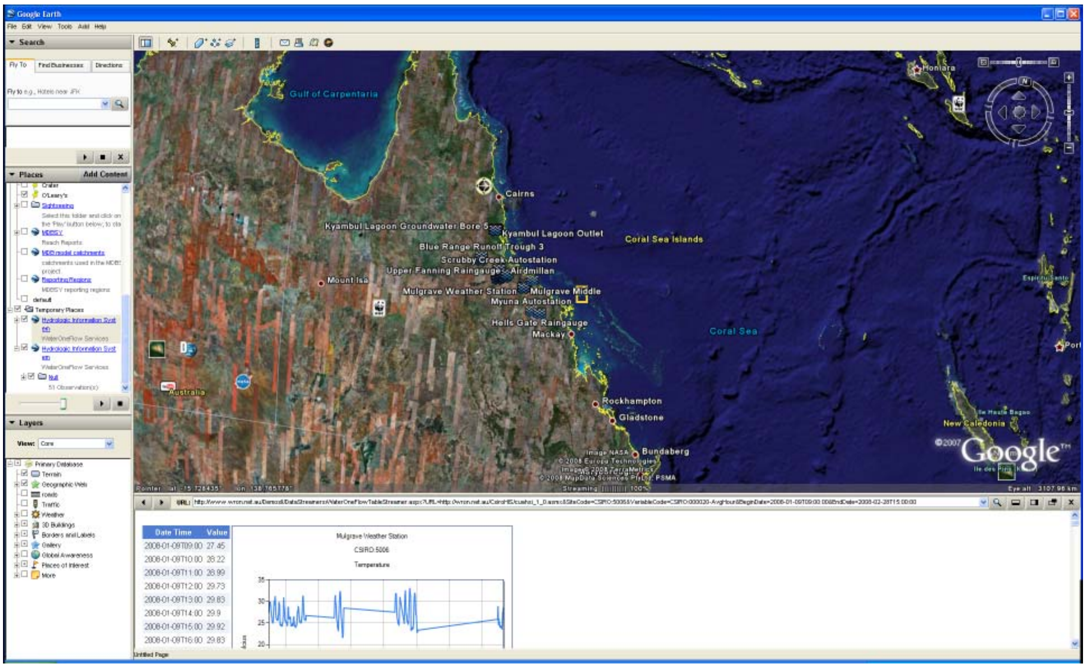

Figure 9 KML output from the CSIRO CUAHSI to KML gateway service

3.4 3D Visualisation

Some hydrological data sets describe inherently three-dimensional structures. For example, digital elevation models (DEM) describe the topography of a region or the 3D shape of a surface; groundwater basins extend into the earth, and so can be described using 3D geometry and boreholes have data that form three dimensional profiles. While it is possible to visualise three-dimensional data in two dimensions utilising colour, shading or sets of profiles, it is often desirable to view objects as three dimensional shapes, or observe how variables change over a three dimensional surface. Additionally, thematic or attribute data visualisation can utilise the third dimension to portray information more clearly over a 2D map image.

There are now several methods of delivering three dimensional content over the internet (note that in the context of this report, 3D content refers to 3D structures or objects that are visualised on a standard monitor – sometimes called 2 ½ D, not objects that the user visualises in 3D through stereo glasses or other techniques, although this is also possible). Techniques include the use of geobrowser tools such as Google Earth or NASA's WorldWind, the development of 3D models in XML-like languages such as VRML or X3D that can be viewed directly within a browser (with an appropriate plugin).

This section provides an overview of key technologies, and of OGC standards and proposals relevant to 3D content delivery.

3.4.1 VRML and X3D

The Virtual Reality Modelling Language (VRML) is a file format for describing interactive 3D objects and worlds. VRML is capable of representing static and animated dynamic 3D and multimedia objects with hyperlinks to other media such as text, sounds, movies, and images (ISO/IEC 14772-1 1997). Delivery of VRML within web browsers is made possible through plugins. Although VRML is still widely used, it is being succeeded by the new Extensible 3D (X3D) standard. X3D is a software standard for defining interactive web- and broadcast-based 3D content integrated with multimedia. X3D is intended for use on a variety of hardware devices and in a broad range of application areas such as engineering and scientific visualization, multimedia presentations, entertainment and educational titles, web pages, and shared virtual worlds. X3D improves upon VRML with new features, advanced application programmer interfaces, additional data encoding formats, stricter conformance, and a componentized architecture that allows for a modular approach to supporting the standard (ISO/IEC 19775 2008).

There are a number of examples of the use of VRML and/or X3D in visualising hydrology and other geoscience data, both on the desktop, and on the web. In (Huang 2004), VRML visualisation is used within a simulation environment to allow interactive steering or tuning of model parameters. The simulation uses TOPMODEL to predict rainfall run-off generation and at each time step, the system generates a 3D visualisation of soil saturation status allowing understanding of watershed responses to rainfall with respect to different terrain features. An example of the visualisation is shown in Figure 10.

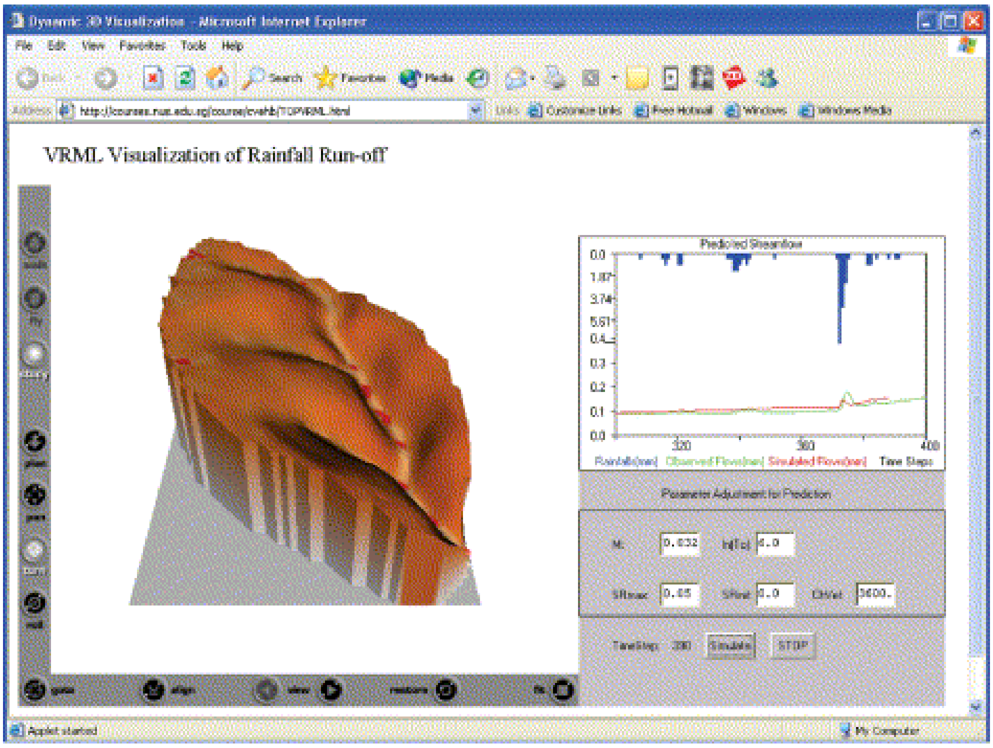

Figure 10 VRML visualisation of saturated soil cells (red areas), source: (Huang 2004)

Ramasundaram et. al. (2005) used interactive VRML models to help students understand the spatial distribution of soils and hydrologic patterns. Three-dimensional models of soil landscapes (soil layers and profiles) and animated 3D models of water tables were developed, with interaction allowing zooming, panning and rotation. The models were then published to a web server for use via web browsers with an installed plugin. Images of the models are shown in Figure 11.

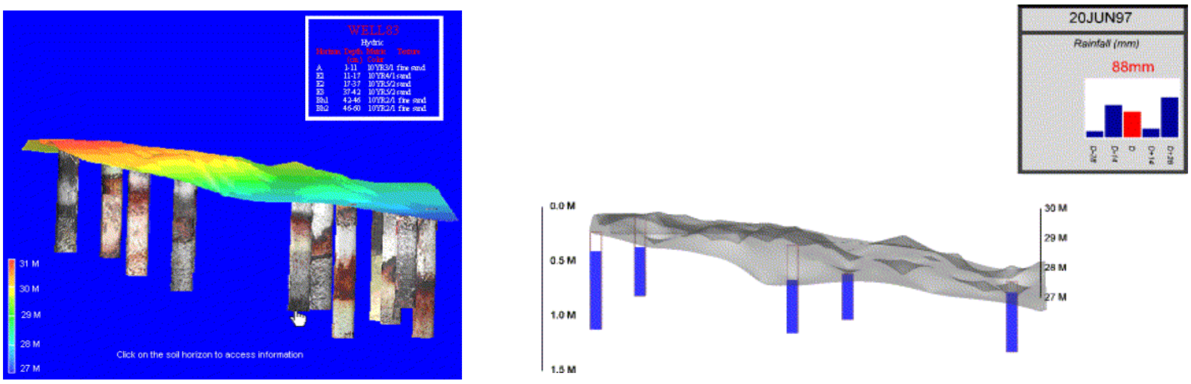

Figure 11 VRML model of soil profile (left) and snapshot from water table dynamics model (right), source: (Ramasundaram 2005)

Geoscience Australia also uses VRML and X3D to develop 3D models of geological systems for the web and to help scientists better understand and interpret their data. Their models are available to the public from http://www.ga.gov.au/map/web3d/. A snapshot of one of Geoscience Australia's models is shown in Figure 12.

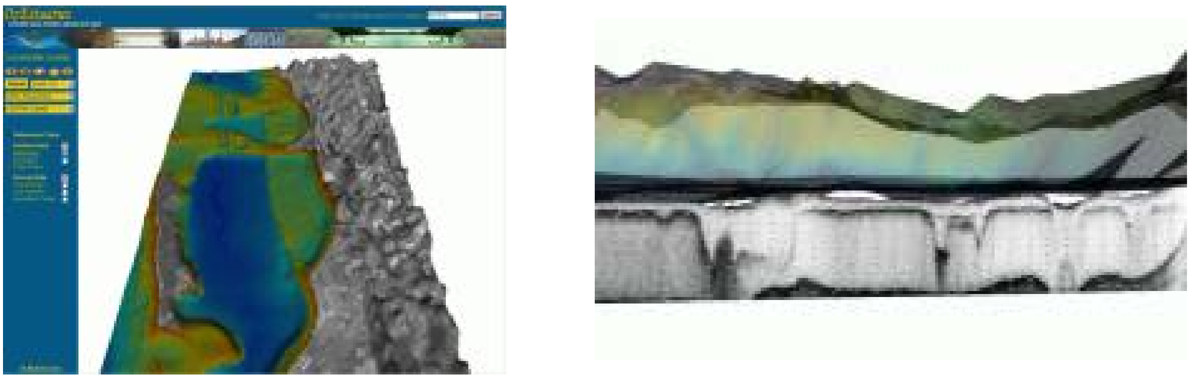

Figure 12 VRML models of Cockburn Sound, source: Geoscience Australia

A number of tools exist to assist in the development and publication of VRML & X3D models. These range from editors, modelling packages that can export in VRML/X3D formats and viewers for both desktop and web use through to post-processing tools and optomisers. In addition, several scientific modelling or visualisation packages can export data in VRML/X3D. These include ESRI's ArcScene, GOCAD and BS GeoFormer. Output from these tools may, however, still require post-processing as it is not always particularly efficient and can result in 3D models with very large file sizes (hundreds of megabytes).

3.4.2 Web Terrain Server

The OGC Web Terrain Server (WTS) discussion paper (Singh 2001) (version 0.3.2, 2001-08-24 available at http://portal.opengeospatial.org/files/?artifact_id=1072) proposes a server for producing perspective views of geo-referenced data. Essentially, this paper proposes an extension to the Web Mapping Service (WMS) defining an extra ***GetView*** operation which specifies the desired view point from which the map image is generated.

An example of a GetView request is:

http://alpha.skylinesoft.com/services/ogc/WMS/WMS3D.asp?request=GetView&srs=EPSG:4326&poi=-84.405049,33.745012,0&distance=500&pitch=90&yaw=0&aov=53&width=512&height=300&format=jpeg&quality=medium

In this example the requested pitch is 90 degrees and yaw is 0 degrees, resulting in what is essentially the same as a GetMap request except that the bounding box is undefined. This example returns the image shown in Figure 13.

Figure 13 WTS GetView example with pitch=90 and yaw=0

The equivalent example with pitch set to 30 degrees would have the following request, with the resulting image shown in Figure 14:

http://alpha.skylinesoft.com/services/ogc/WMS/WMS3D.asp?request=GetView&srs=EPSG:4326&poi=-84.405049,33.745012,0&distance=500&pitch=30&yaw=0&aov=53&width=512&height=300&format=jpeg&quality=medium

Figure 14 WTS GetView with pitch=30

The parameters that define the view are shown in Figure 15.

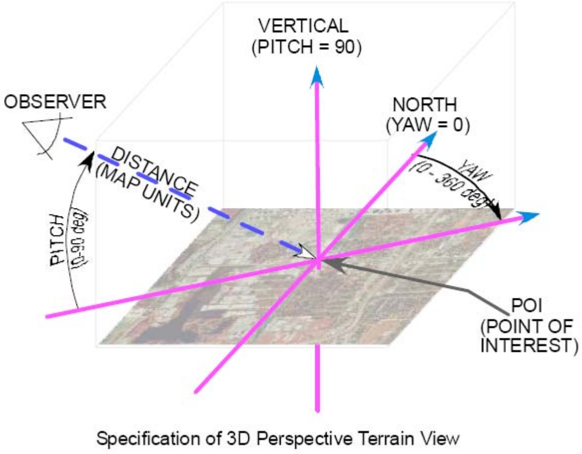

Figure 15 WTS view parameters

Example implementations of the Web Terrain Server include *terrainServer* (http://www.conterra.de/en/products/sdi/terrainserver/index.shtm) which claims to fully support the WTS interface. This product actually produces 3D views by draping 2D image data from any map service as a texture on a terrain model. An online demonstration is available at http://www.sdi-suite.de/en/terrain-server.shtm. The OGC implementations page also lists other products that claim to support WTS.

3.4.3 Web 3D Service (W3DS)

Another OGC discussion paper, the Web 3D Service (Quadt & Kolbe 2005) (version 0.3.0, 2005-02-02 available at http://portal.opengeospatial.org/files/?artifact_id=8869), specifies a portrayal service for three-dimensional geodata, delivering graphical elements from a given geographical area. In contrast to the Web Mapping Service (WMS) and the Web Terrain Server (WTS) which produce map-like images, the W3DS produces 3D scene graphs. A scene graph is an aggregation of 3D graphical elements, which can be explored interactively by a user. The W3DS merges different types (layers) of 3D data in one scene graph. VRML97 is the default output format, however other more recent formats like GeoVRML or X3D can also be used. In the W3DS the term 'layers' is used as a synonym for sets of different graphical 3D elements.

The Web 3D Service provides two operations – GetCapabilities, and GetScene. ***GetCapabilities*** returns an XML document describing the W3DS implementation's capabilities, including what 3D formats it can produce and what layers are available. ***GetScene*** returns a 3D scene graph. The GetScene request includes parameters that define the virtual camera position (or view) such as point of interest, distance, pitch, yaw and angle of view as well as layers, bounding box and desired output format parameters. An example GetScene request is as follows:

http://www.myserver.de/W3DS?VERSION=0.3.0&REQUEST=GetScene&SRS=EPSG:31466&FORMAT=model/vrml&POI=2583150.0,5720600.0,52.5&PITCH=45&YAW=45&DISTANCE=100&AOV=90&BBOX=2583150,5720600,2584000,5722000&LAYERS=dtm,buildings,vegetation&STYLES=orthophoto,,simple

Demonstration implementations of the W3DS do not seem to be available currently, although the University of Bonn GDI3D project (http://www.geographie.uni-bonn.de/karto/hd3d/index.en.htm) implementation is discussed in (Zipf 2007). This implementation seems to be targeted largely at delivering 3D city-scapes and urban models. A related standard for this type of content is CityGML (http://www.citygmlwiki.org/index.php/Main_Page).

3.5 Visualisation of Temporal Data

Visualisation of temporal and spatio-temporal data is an important aspect of AWRIS. Observational data sets such as precipitation, dam levels, stream flow or water quality measures are all time-series data. Even feature data sets such as gauging stations or monitoring points have a temporal component when considered as a set of features whose existence, identity or attributes change over time. Consideration should therefore be given to appropriate ways to present temporal information over the web.

Temporal information is inherently associated with change over time. In a spatial context, changes may be characterised as (Andrienko 2003):

1. Existential changes, i.e. appearance and disappearance.

2. Changes of spatial properties: location, shape or/and size, orientation, altitude, height, gradient and volume.

3. Changes of thematic properties expressed through values of attributes: qualitative changes and changes of ordinal or numeric characteristics (increase and decrease).

Different presentation techniques are suited to particular types of changing data. In this section, several general techniques for spatio-temporal presentation are described, including information on the types of changes they are useful at representing.

3.5.1 Map Iteration

Map iteration is a traditional approach for presenting information on change. Change is visualised by showing images or maps at different stages of change side by side or in a sequence so that changes can be seen and compared. Map iteration is a general presentation technique, as existential, spatial and thematic changes can all be portrayed. This technique can be useful for presenting temporal changes, but can also be used to compare different result sets from, for example, a set of scenarios. The obvious disadvantage of this technique is that a limited number of images can be displayed simultaneously, thus limiting the resolution of change that can be shown.

3.5.2 Animation

Animation of web components exploits the ability of computers to deliver dynamic content. In contrast with the static images of paper reports and documents, the web allows for dynamic, interactive maps, charts and images to be presented. This is especially useful when presenting temporal information although animation can also be used for non-temporal information where time is used to reveal connected features and put them in a logic sequence (Oberholzer 2000, Dransch 1995).

Animation can be used to show changes in spatial or thematic attributes of objects. If the location, size or shape of a feature changes over time, this can be represented either directly in the images or maps, or indirectly through symbology changes. Thematic changes (such as changing water quality measures) can be shown using symbols, colours or movement.

Ideally, components that provide animated content should also provide control over the animation. In particular, the following features are useful to control, either during creation of the animation, or by the user (adapted from (Andrienko 2003)):

- Speed of the animation;
- Direction – forth or back;
- Extent (start and finish points);
- Moments/intervals to include in animation:
  - Step (i.e. the interval between time moments successive animation frames refer to);
  - Moments or periods within a cycle;
  - Arbitrary selection.
- Smoothness (including interpolation between frames).

Fading and morphing are sometimes used in association with animation to show changes progressively in a sequence of images so that changing elements may be visible in more than one image at a time.

Animation can be useful in portraying changes including existential changes, spatial changes and thematic change. A detailed analysis of data is, however, difficult with animations, and periodic or cyclical trends can be difficult to identify.

3.5.3 Linked Views

Often one particular view of data (for example a 2D map view) is insufficient to allow proper understanding of a data set. Linked views refer to interfaces with several windows displaying different perspectives, dimensions, or thematic parameters of a data set, all of which are dynamically linked. This might provide, for example, a navigable 2D spatial representation with thematic parameters represented in colour, and an associated profile view showing sub-surface structures, both linked in the temporal dimension so that moving a time-slider changes the state of both views simultaneously.

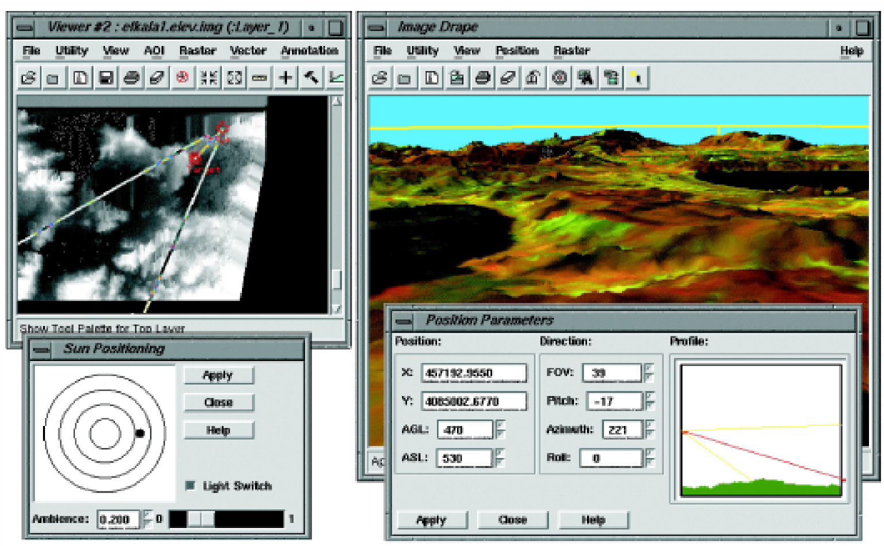

Figure 16 A set of dynamically linked views, source: (Uhlenkuken 2000)

Timelines are a commonly linked component – a timeline is displayed along with the map view and current time indicated as the thematic and spatial content changes on the map.

3.5.4 Time in the Third Dimension

The use of the third dimension as a temporal axis allows thematic changes to be displayed over particular locations. In these examples, a 2D map is displayed, and the third (vertical) dimension is used to display details of parameters over time. Variations in colour or size of indicators can represent the characteristics of events or parameter changes. Figure 17 shows an example of this ‘space-time cube’ approach where events are displayed by circles in the vertical axis. A similar approach is given in (Tominski 2005) and shown in Figure 18. In this instance, multiple attributes are displayed at once using each side of the pencil (left image) for a different parameter. In the right image, cyclical patterns in the data are emphasised through displays using helix icons.

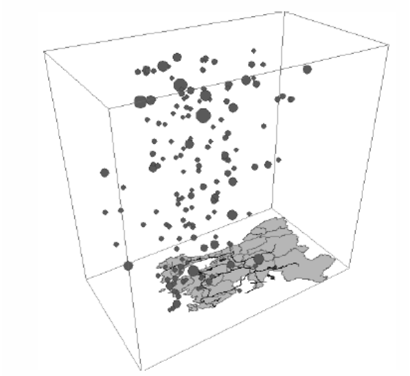

Figure 17 "Space-Time Cube" showing events at different times (vertical axis), source: (Gatalsky 2004)

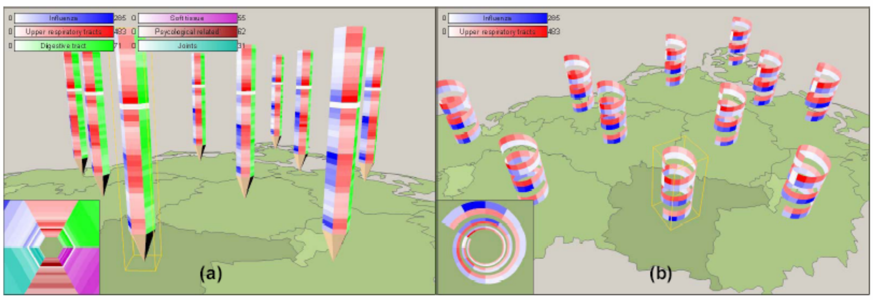

Figure 18 Temporal changes using the vertical axis and pencil (left) or helix (right) icons, source: (Tominski 2005)

3.5.5 Embedded Graphs

Embedding graphs and other diagrams within map images can also be useful in displaying temporal aspects of a data set. This is particularly common in applications such as GoogleMaps where extra content can be linked to marked sites in the map display. Selecting a site can then display a time-series graph (see Figure 19). A disadvantage of this approach is that the graph can temporarily obscure significant regions of the map, and graphs from multiple sites can not be compared.

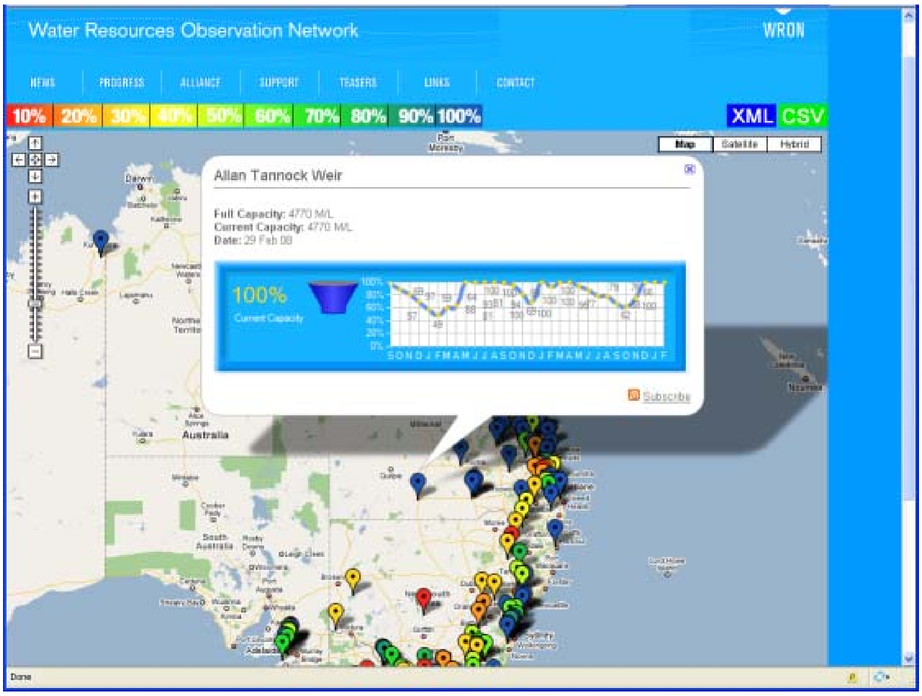

Figure 19 GoogleMaps interface showing an embedded graph, source: http://www.wron.net.au/DemosII/DamData/DamMap.aspx

3.6 Analysis

While visual presentation of water information will meet the needs of some users, the ability to interactively analyse available data enables real value-added insight to be obtained. Data analysis can refer to a variety of techniques from graph or chart generation, to simple statistical operations such as averaging values, or complex processes applied to data sets to generate new information. This section provides an overview of several analysis techniques, and the web technologies that support their implementation.

3.6.1 Graphing

Graphs are used extensively in current water account and assessment reports, particularly as a means of portraying trends. While the current use of graphs and charts is likely to be largely driven by their suitability for the current medium of report delivery (i.e. printed documents), there will still be a need for graphs and charts for data analysis within an online system.

An evaluation of the State Water Report for Victoria 04/05 and 05/06 (Victorian Government 2006, Victorian Government 2007) provides an indication of the types of water information that might be graphed, and the types of graphs used. The results are summarised in Table 1 below.

Variable	Graph Type	Description
Stream flow	Bar chart	Annual stream flow values since 1960 for a particular site

Water storage levels	Line chart	Comparison of monthly storage levels for the last ten years for rural or city storages
	Bar chart	Monthly storage levels as percentage full for last ten years (rural and city storages)
Consumptive use	Pie chart	Proportion of consumption for various uses (e.g. irrigation, urban, rural)
Water restrictions	Line chart	Number of towns on water restrictions each month for two years
	Line chart	Number of unregulated streams on restrictions each month for two years
Runoff	Box plot	Predicted percentage change in runoff per basin

Table 1 Types of graphed data in State Water Report for Victoria 04/05

In addition to the items in Table 1 above, the State Water Report for Victoria also provides data on a large number of other water resource indicators in the form of tables and maps. The information in these tables may also be suited to graph delivery, particularly with the added capabilities of web graphing tools.

As well as production of static image graphs, web technologies also allow for more interactive graph delivery. Flexibility can be greatly increased by dynamic generation of graphs in response to particular user queries, rather than provision of pre-prepared analyses only. In this way, support for a wider variety of user needs can be accommodated without the need to fully understand or prepare for all possible user requests. In the case of the graphs in Table 1, this means that the user could select the range of years for which they want to see water storage levels, or identify a set of sites to compare stream flow for.

Graphing toolkits also allow greater interaction with the graph itself, including the ability to zoom in or out to view the data at higher or lower resolution, or the ability to animate data over a time series. This allows for more detailed analysis and potentially investigation of multiple variables over time. As an example, compare the following graphs of stream flow. The first graph (Figure 20) is from the State Water Report for Victoria 04/05 and shows annual stream flow at Wimmera River. The remaining graphs (Figure 21) show daily stream flow for Swan Hill (source data from MDBC) from 1978 to 2006 at different time scales. These graphs are snapshots from an interactive graph (accessible at http://charger-bt.act.cmis.csiro.au:8080/chart/Chart.html).

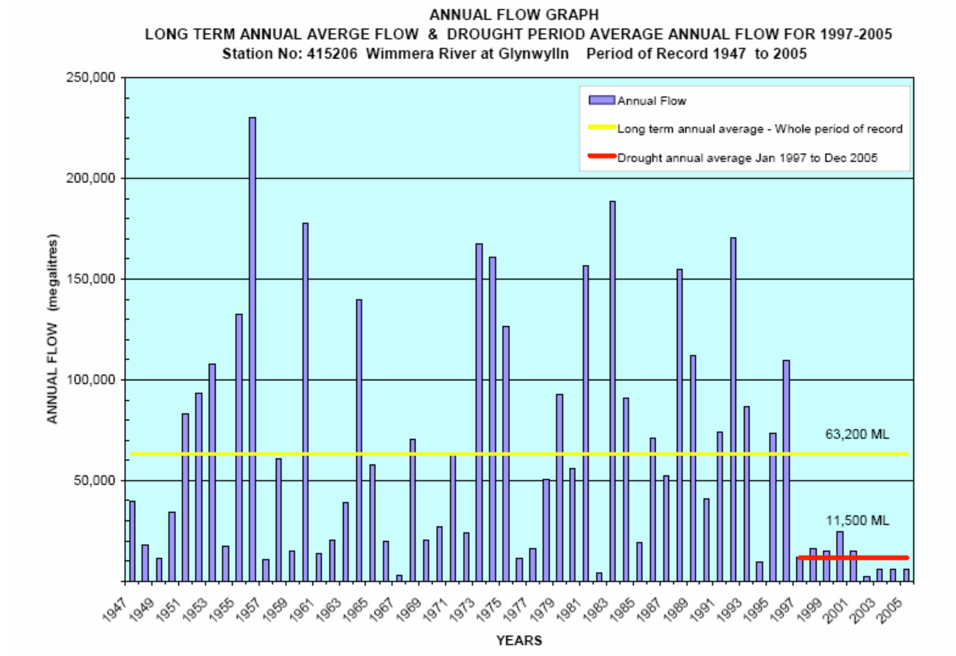

Figure 20 Annual stream flow at Wimmera River, source: State Water Report for Victoria 04/05

It is also useful to consider graphing within a spatial context. While graphs displaying trends and values for multiple sites are useful, they lack the ability to portray spatial relationships between the sites at the same time. The combination of graphs with maps or spatial displays may enrich this type of information presentation as discussed in Section 3.5.5 above.

Delivery of graphs over the web can be achieved in one of two ways:

1. serve graph images from a directory on the web server; or

2. dynamically generate the graph and stream the object to the client browser.

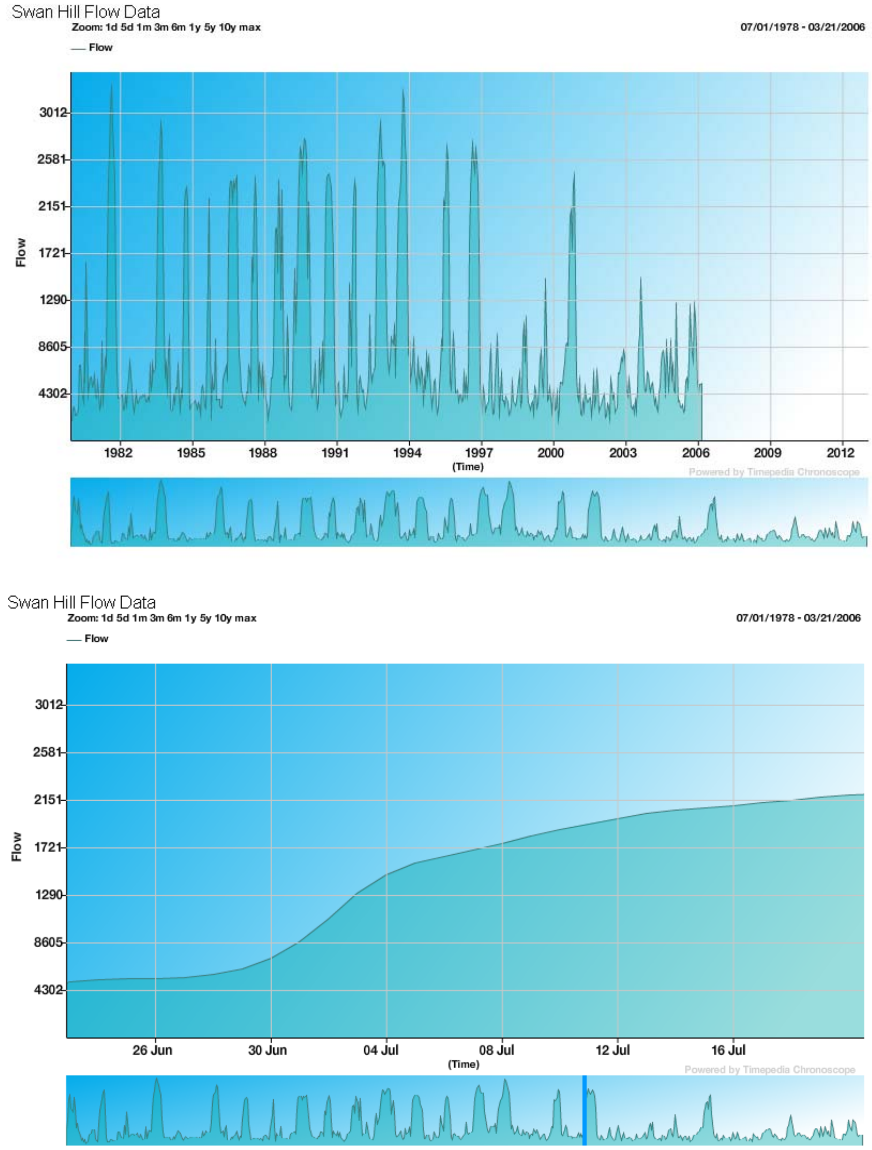

Figure 21 Swan Hill stream flow data, entire dataset (top) and zoomed in to show daily values (bottom)

Serving graphs from a web server directory

This involves pre-generating specific graph images, storing the images on a file system and allowing the client browser to retrieve the file via a web server. This has the advantage of having minimal programming overhead, however this approach requires a clean-up cycle to ensure images are removed appropriately so that excessive storage is not required. Access is generally slower than database access due to file system overheads. In a web farm environment files generally need to be distributed and maintained synchronously on all machines within the farm. This approach is limited in terms of scalability of content. As an example, consider a data set with PH, turbidity and temperature variables for a number of sites. Each site could potentially result in up to seven graphs (including variable combinations). It would swiftly become unmanageable and unscalable for large numbers of sites if graphs were pre-generated.

Dynamic graph generation & binary streaming of objects

In this approach, a web server application generates a graph on the fly in response to each individual user request and the resulting image object is returned directly to the client browser as a binary stream. This approach provides much greater flexibility as many different graphs can be produced dynamically rather than only providing a limited set of pre-generated graphs. It also eliminates the need for multiple copies of a file in a web farm environment, and for file system storage of image files. Serving graph images in this way requires more direct access to the underlying data sources as the graphs are computed in response to the request.

In both the above cases, caching can be used to reduce the load where multiple users request the same graph. This is generally handled by the proxy server.

As with other visualisation and reporting components, the presentation layer should be kept separate from the data layer wherever possible.

3.6.2 Data cubing

Allowing analysis of very large data sets presents significant problems for web-based systems. Data sets may contain billions of rows of data, and providing summary information or relational queries across these data sets requires time-consuming processing resulting in unacceptable delays in producing results. As an example, consider a region such as the Border Rivers. If a user wants to retrieve annual runoff from 1895 – 2006, spatially averaged across the region, some 100 million rows of data may need to be aggregated. This computation would take in the order of several minutes to complete on a current, high-end database server – a response time unacceptable for website users. Data cubes provide a solution to this problem by pre-aggregating data into multi-dimensional tables that are optimised for particular types of queries. Asking the same question of a configured data cube would return values in sub-second response times.

Data cubes, or OLAP (OnLine Analytical Processing) cubes, are multi-dimensional extensions of 2-D tables, just as in geometry a cube is a three-dimensional extension of a square. 3-D data cubes can be thought of as a set of similarly structured 2-D tables stacked on top of one another (Kay 2004). At the intersection points of a cube's dimensions, analysts can view the measures. Measures are numeric values, usually business metrics (such as sales, profit and cost of goods), at a given set of dimension intersections. For example, to view the sales of a given product at a given store at a given time, examine the cube at the point where those dimensions intersect to find the measures (IBM). OLAP performs multidimensional analysis of data and provides the capability for complex calculations, trend analysis, and sophisticated data modelling. OLAP enables end-users to perform ad hoc analysis of data in multiple dimensions, thereby providing the insight and understanding they need for better decision making and targeting data bounds that require further data investigation.

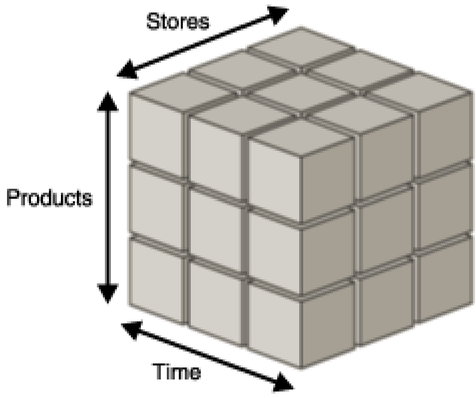

Figure 22 Example of dimensions of a data cube (source: IBM DB2 Alphablox Cube Server Overview)

There are three different types of OLAP as discussed below. Each of the types is optimised for a particular aspect – query return, complex calculations, data size, and has corresponding disadvantages. The choice of OLAP type will depend on the requirements of the system.

MOLAP – Multidimensional Online Analytical Processing

This is the traditional form of OLAP analysis. In MOLAP, data is stored in a multidimensional cube. The storage is not in the relational database, but in proprietary formats (Rojas 2006).. The advantages of MOLAP are excellent performance – MOLAP cubes are built for fast data retrieval, and are optimised for slicing and dicing operations (Rojas 2006); and efficient performance of complex calculations – since all calculations are pre-generated when the cube is created, complex calculations are not only feasible, but they return quickly. The main disadvantage of MOLAP is that it is limited in the amount of data it can handle. Because all calculations are performed when the cube is built, it is not possible to include a large amount of data in the cube itself. This is not to say that the data in the cube cannot be derived from a large amount of data, but in this case, only summary-level information will be included in the cube itself (Rojas 2006). Additionally, as the number of dimensions increases, the cube becomes sparser – that is, many cells representing specific attribute combinations are empty, containing no aggregated data (Kay 2004). Finally, implementing MOLAP requires additional investment as cube technologies are often proprietary and do not already exist in the organisation. Therefore, to adopt MOLAP technology, chances are additional investments in human and capital resources are needed (Rojas 2006).

ROLAP – Relational Online Analytical Processing

This methodology works directly with relational databases and relies on manipulating the data stored in the relational database to give the appearance of traditional OLAP functionality without the need to pre-compute and store information. Instead, SQL queries are generated dynamically to calculate information at the appropriate level in response to user requests (Rojas 2006, Wikipedia – ROLAP). The main advantage of ROLAP is that it can handle large amounts of data – the data size limitation of ROLAP technology is the limitation on data size of the underlying relational database. In other words, ROLAP itself places no limitation on data amount (Rojas 2006). ROLAP can also leverage functionalities inherent in the relational database such as row-level security and filter results based on preset criteria for specific groups of users (Rojas 2006, Wikipedia – ROLAP). The disadvantage of ROLAP is that performance can be slow. Because each ROLAP report is essentially a SQL query (or multiple SQL queries) in the relational database, the query time can be long if the underlying data size is large (Rojas 2006). ROLAP is also limited by SQL functionalities. Because ROLAP technology mainly relies on generating SQL statements to query the relational database, and SQL statements do not fit all needs (for example, it is difficult to perform complex calculations using SQL), ROLAP technologies are therefore traditionally limited by what SQL can do. ROLAP vendors have mitigated this risk by building into the tool out-of-the-box complex functions as well as the ability to allow users to define their own functions (Rojas 2006).

HOLAP – Hybrid Online Analytical Processing

HOLAP technologies attempt to combine the advantages of MOLAP and ROLAP by storing part of the data in a MOLAP store and another part of the data in a ROLAP store. Data partitioning can occur by storing summary-type information (aggregations) in MOLAP for fast query performance and detailed data in ROLAP to optimise processing. Alternatively, recently requested slices may be stored in MOLAP for fast query performance and older data in ROLAP (Rojas 2006, Wikipedia – HOLAP).

4. RECOMMENDATIONS

4.1 General Observations

In evaluating all the component areas in this report, several general observations can be made:

- Detailed user requirements analysis is needed;
- The spatio-temporal nature of water resources information needs to be considered; and
- Further analysis of specific AWRIS usage scenarios is needed.

While the AWRIS User Requirements document (NWC 2006) details some of the desired features for a data explorer and data dashboard, specific details on what particular user groups would use the system for, and why, remain unclear. This is at least partially due to the difficulty of determining how a currently non-existent system might be used. However, an understanding of why users will use the system and what they will use it for, will be key in determining how the system should be developed, and critical to how usable and effective the resulting system becomes. The design process for AWRIS web applications should include a detailed analysis of the requirements of potential user groups. This requirements analysis needs to capture the primary use scenarios – that is, what questions will users want answered by the system – in order to determine the best approach to interface and tool design. These use scenarios will underpin development of all the components mentioned in Section 3 – the query interface, how map displays can be best used, what temporal information will need to be portrayed and how, and how best to utilise technologies such as data cubing to optimise data retrieval.

The spatio-temporal nature of water resource information should also be considered in the design of AWRIS web applications. Water information is inherently associated with location, and the spatial relationships between features are important. The ability of the system to support presentation and discovery of spatial information will greatly impact on it's usefulness to users. In addition, since observational data has a strong temporal component and will form a large proportion of the total data holdings, interfaces and visualisations need to be developed within a combined spatio-temporal context.

Finally, following on from the user requirements analysis suggesting in the first point, an in-depth investigation of presentation techniques should be carried out based on the key questions and use scenarios. This would include evaluation and prototyping of various ways to portray the selected information sets in order to address specific queries. This targeted approach for specific water information scenarios would enable a more effective design of the system in the long term. An exploration of techniques for effectively displaying combinations of spatial and temporal data would also assist in the design process.

4.2 Non-functional Requirement Considerations

In addition to the general observations noted in Section 4.1, there are a number of non-functional requirements that need to be considered. These involve the organisational environment in which AWRIS is to be developed and deployed. Areas of particular concern are:

- User experience, particularly in terms of system performance;
- Bandwidth limitations; and
- Organisational constraints on content delivery methods.

System performance will be a key issue in the development of AWRIS. The ability of the system to respond to user queries in ‘web time’ will directly influence the user’s experience. Given the large volume of data that will become available, techniques for efficient data management are crucial to the system’s success. Technologies such as OLAP can assist in this area, particularly where data cube dimensions are developed in conjunction with primary use scenarios. System architecture can also have a large impact on performance.

Limitations on bandwidth imposed by the user community will have a significant impact on AWRIS. While research and government departments may have access to high-speed internet, it can be assumed that a large proportion of the general public, particularly rural users, do not. AWRIS web applications will need to be designed with this in mind. Research into suitable formats for data transfer, and particularly binary implementations of XML formats like KML will be necessary as the volume of data transmitted, even for visualisation, becomes greater. Alternative approaches involving server-side rendering, where this is efficient, should also be explored.

The third area for consideration is the organisational development and deployment environment. It is possible that constraints exist on the types of content that AWRIS web applications can present, and these constraints need to be clearly articulated. For example, if only static content can legally be provided, the variety of presentation methods is significantly reduced and the flexibility and usability of the system will be restricted. While beyond the scope of this document, eliciting these constraints will be a necessity of the requirements capture process.

This section is not intended to provide an exhaustive list of the issues involved in developing AWRIS, but rather to serve as a starting point for areas that need to be investigated further with respect to data discovery, portrayal and analysis.

REFERENCES

Andrienko, N., Andrienko, G. & Gatalsky, P. 2003 'Exploratory spatio-temporal visualization: an analytical review', *Journal of Visual Languages and Computing*, vol 14, no. 6, pp. 503-541.

AWRIS 2005, *Australian Water Resources Information System*, available at http://www.water.gov.au/Keymessages/AustralianWaterResourcesInformationSystem/index.aspx?Menu=Level1_1_13

Deb, B., Bannur, S.G. & Bharti, S. 2007, 'Rich Internet Applications (RIA) – Opportunities and Challenges for Enterprises', *InfoSys White Paper*, available at http://www.infy.co.uk/technology/rich-internet-applications.pdf

de la Beaujardiere, J. (ed.) 2004, *OGC Web Map Service Interface*, Open GIS Consortium Inc. project document OGC 03-109r1.

Dransch, D. 1995, *Temporale und nontemporale Computer-Animation in der KartographieBerliner Geowissenschaftliche Abhandlungen, Reihe C Kartographie* **15** Fachbereich Geowissenschaften, FU Berlin, Berlin, Germany.

Garrett, J. J. 2005, *AJAX: A New Approach to Web Applications*, available at http://www.adaptivepath.com/ideas/essays/archives/000385.php.

Gatalsky, P., Andrienko, N. & Andrienko, G. 2004, 'Interactive Analysis of Event Data Using Space-Time Cube', *Proceedings of Eigth International Conference on Information Visualisation*, IEEE Computer Society, Washington D.C., USA, pp. 145-152.

Henry, L. 2007, 'Composers Workbench: A Review of Rich Internet Applications Technologies', *CSIRO Internal Report 07/385*.

Huang, B. & Claramunt, C. 2004, 'Environmental simulation within a virtual environment', *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 59, nos. 1-2, pp. 73-84.

IBM, *DB2 Alphablox Cube Server Administrator's Guide*, available at http://publib.boulder.ibm.com/infocenter/ablxhelp/v8r4m0/index.jsp?topic=/com.ibm.db2.abx.cub.doc/abx-c-cube-cubingreldata.html.

ISO/IEC 14277-1 1997, *Information technology - Computer graphics and image processing - The Virtual Reality Modeling Language - Part 1: Functional specification and UTF-8 encoding*, International Organisation for Standardisation, Geneva, Switzerland.

ISO/IEC 19775-1 2004, *Information technology - Computer graphics and image processing - Extensible 3D (X3D) - Part 1: Architecture and base components* , International Organisation for Standardisation, Geneva, Switzerland.

Kay, R. 2004, 'QuickStudy: Data Cubes', *ComputerWorld Business Intelligence*, available at http://www.computerworld.com/action/article.do?command=viewArticleBasic&articleId=91640.

LaMar, E. (ed.) 2005, *OGC Web Map Service – Proposed Animation Service Extension*, Open Geospatial Consortium Inc. project document OGC 06-045r1.

Lupp, M. (ed.) 2007, *Styled Layer Descriptor profile of the Web Map Service Implementation Specification*, Open Geospatial Consortium Inc. project document OGC 05-078r4.

NWC 2006, *User Requirements for the Australian Water Resources Information System*, National Water Commission, Canberra: Commonwealth of Australia.

Oberholzer, C. & Hurni, L. 2000, 'Visualization of change in the Interactive Multimedia Atlas of Switzerland', *Computers & Geosciences*, vol. 26, no. 1, pp. 37-43.

Peuquet, D. J. 1994, 'It's about time: a conceptual framework for the representation of temporal dynamics in geographical information systems', *Annals of the Association of American Geographers*, vol. 84, no. 3, pp. 441-461.

Quadt, U. & Kolbe, T. H. (eds) 2005, *Web 3D Service*, Open Geospatial Consortium Inc. project document OGC 05-019.

Ramasundaram, V., Grunwald, S., Mangeot, A., Comerford, N. B. & Bliss, C. M. 2005, 'Development of an environmental virtual field laboratory', *Computers and Education*, vol. 45, no. 1, pp. 21-34.

Rojas, E. 2006, 'DataMart (Data Warehouse) Tool:Mondrian + JRubik', (presentation) *International Crop Information System (ICIS) Workshop*, accessed 10 Feb 2008, http://cropwiki.irri.org/icis/images/e/ea/DataMart-Tool.pdf.

Singh, R. R. (ed.) 2001, *OGC Web Terrain Server (WTS)*, Open GIS Consortium Inc. project document OGC 01-061.

Stewart, B. 2008, WIRADA Data Delivery, Project Agreement No. 30/4612/2, Bureau of Meteorology internal document ref. number 35/451.

Tominski, C., Schulze-Wollgast, P. & Schumann, H. 2005, '3D Information visualization for time dependent data on maps', *Proceedings of Ninth International Conference on Information Visualisation*, IEEE Computer Society, Washington D.C., USA, pp. 175-181.

Uhlenkuken, C., Schmidt, B. & Streit, U. 2000, 'Visual exploration of high-dimensional spatial data: requirements and deficits', *Computers and Geosciences*, vol. 26, no. 1, pp. 77-85.

Victorian Government 2006, 'State Water Report 2004/05: a statement of Victorian water resources', published by the Victorian Government Department of Sustainability and Environment.

Victorian Government 2007, 'State Water Report 2005/06: a statement of Victorian water resources', published by the Victorian Government Department of Sustainability and Environment.

White, A., 'Web RIA Crystal Ball: Where Are We Now and Where Are We Going?', available at http://www.developer.com/services/article.php/3582021.

Wikipedia – ROLAP, 'ROLAP', *Wikipedia, the Free Encyclopedia*, available at http://en.wikipedia.org/wiki/ROLAP.

Wikipedia – HOLAP, 'HOLAP', *Wikipedia, the Free Encyclopedia*, available at http://en.wikipedia.org/wiki/HOLAP.

Wilson, T. (ed.) 2007, *KML 2.2 – An OGC Best Practice*, Open Geospatial Consortium Inc. project document OGC 07-113r1.

Zipf, A., Basanow, J., Neis, P., Neubauer, S. & Schilling, A. 2007, 'Towards 3D Spatial Data Infrastructures (3D-SDI) based on Open Standards - experiences, results and future issues', *"3D GeoInfo07" ISPRS WG IV/8 International Workshop on 3D Geo-Information: Requirements, Acquisition, Modelling, Analysis, Visualisation*, Delft, Netherlands.

Contact Us

Phone: 1300 363 400
+61 3 9545 2176

Email: enquiries@csiro.au

Web: www.csiro.au

Your CSIRO