

# The Australian Hydrological Geospatial Fabric – Development Methodology and Conceptual Architecture

Atkinson, R., Power, R., Lemon, D. and O'Hagan, R.

28 April 2008

Australian Government: Bureau of Meteorology  
Ms Dovey Dee  
Dr David Kinny

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

Enquiries should be addressed to:

Dr David Lemon

David.Lemon@csiro.au

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](http://www.csiro.au/org/HealthyCountry.html)

Citation: Atkinson, R., Power, R., Lemon, D. and O'Hagan, R., Dovey, D. and Kinny, D. 2008. CSIRO: Water for a Healthy Country National Research Flagship, Canberra, Australia, 57 pp.

## **Copyright and Disclaimer**

© 2008 CSIRO To the extent permitted by law, all rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

## **Important Disclaimer**

CSIRO advises that the information contained in this publication comprises general statements based on scientific research. The reader is advised and needs to be aware that such information may be incomplete or unable to be used in any specific situation. No reliance or actions must therefore be made on that information without seeking prior expert professional, scientific and technical advice. To the extent permitted by law, CSIRO (including its employees and consultants) excludes all liability to any person for any consequences, including but not limited to all losses, damages, costs, expenses and any other compensation, arising directly or indirectly from using this publication (in part or in whole) and any information or material contained in it.

## **ACKNOWLEDGEMENTS**

The authors wish to thank the Bureau of Meteorology for their funding of this project through the WIRADA Bridging Projects and especially the support and guidance of Dr David Kinny and Dovey Dee.

The authors would also like to thank the many people who provided input into this project including:

- Amit Parashar (CSIRO ICT Centre)
- Simon Cox (CSIRO Exploration and Mining)
- Peter Fitch (CSIRO Land and Water)
- Albert van Dijk (CSIRO Land and Water)
- Cuan Petheram (CSIRO Land and Water)
- Russell Crosbie (CSIRO Land and Water)
- Santosh Aryal (CSIRO Land and Water)
- John Gallant (CSIRO Land and Water)
- Arthur Read (CSIRO Land and Water)
- Graham Turner (CSIRO Sustainable Ecosystems)

## EXECUTIVE SUMMARY

The Australian Hydrological Geospatial Fabric (AHGF) is a suite of well maintained, evolving, authoritative, data products containing a consistent representation of every hydrologic feature, and the connectivity between features, of the Australian water system. These products are capable of supporting a wide variety of uses for many users. As such, the AHGF will become the framework geospatial information upon which Australia's water information related activities are based and through which they are related.

For the initial implementation of the AHGF these baseline data products will need to provide national coverage of the following data sets:

- 1- Surface Water – Hydrology;
- 2- Surface Water – Measurement Sites;
- 3- Groundwater – Hydrology;
- 4- Groundwater – Measurement Sites;
- 5- Digital Elevation Model (DEM) at 1 arc-second resolution; and
- 6- Key Catchments.

Implementing the AHGF will require consideration of a number of architectural issues not commonly addressed when building similar products. Key amongst these is the expectation of a need for the AHGF to support a large number of use cases, many of which remain undocumented. This requires that the AHGF is built around a flexible, evolvable information model.

A review of similar initiatives has found that current practice in development of hydrologically focussed products results in compromised flexibility of the final data product, primarily due to constraints arising from the implementation. This means that such products do not support the number and variety of use cases that the AHGF may need to support.

As understanding of the nature and requirements of the AHGF has developed, it has become clear that the keys to success will be:

- a modular conceptual information model, capable of supporting many use cases and associated implementations. This will be the foundation on which implementations will be developed;
- consistent use of standardised Data Product Specifications describing clearly both data sets input to the AHGF and derived data products output from the AHGF; and
- adoption of a nimble, iterative, requirements-focussed development methodology that covers all aspects of the data product lifecycle and allows for evolution of technology, data products and user requirements.

This report focuses on developing and describing a methodology for building the AHGF. This methodology is designed around the critical success factors described above. Starting points for the specification of components of the AHGF are proposed. A conceptual data delivery

architecture and elements of the conceptual information model for surface water hydrology are given. These need to be reviewed in the context of critical use cases.

While this report emphasises the importance of developing a sound conceptual model, the process of implementation of the AHGF should be a parallel task during the design phase, helping to expose issues and requirements and resulting in a more robust conceptual model as the foundation for an evolvable system.

# Contents

- 1. Introduction ..... 1**
  - 1.1 Background..... 1
  - 1.2 The Australian Hydrological Geospatial Fabric .....2
  - 1.3 Use of the AHGF.....3
  - 1.4 Project History.....4
  - 1.5 Report Structure.....4
  
- 2. Review of Hydrological Geospatial Fabrics ..... 5**
  - 2.1 US Hydrological Reference Data Sets.....5
    - 2.1.1 National Hydrography Data set (NHD) ..... 6
    - 2.1.2 NHD Plus ..... 7
  - 2.2 European Activities ..... 9
    - 2.2.1 CCM 2..... 9
    - 2.2.2 SANDRE ..... 10
  - 2.3 Other Initiatives ..... 10
    - 2.3.1 Arc Hydro ..... 10
    - 2.3.2 Arc Hydro Groundwater Data Model ..... 11
    - 2.3.3 Canadian Groundwater Data Model ..... 12
  - 2.4 Discussion..... 13
  
- 3. Methodology ..... 15**
  - 3.1 AHGF Development Principles ..... 15
    - 3.1.1 Risks ..... 16
    - 3.1.2 Staged Approach to Data Set Development..... 16
    - 3.1.3 Iterative Development and Rapid Prototyping ..... 17
    - 3.1.4 Separation of Conceptual and Implementation Models ..... 17
    - 3.1.5 End-User Involvement..... 17
  - 3.2 Proposed Methodology ..... 18
    - 3.2.1 Design phases ..... 20
    - 3.2.2 Implementation..... 22
    - 3.2.3 Data maintenance environment testing..... 23
    - 3.2.4 Service interfaces..... 24
  
- 4. Developing a Conceptual Model..... 25**
  - 4.1 Conceptual and Implementation Models ..... 25
  - 4.2 Spatial modelling..... 27

4.2.1	Geometry Modelling .....	27
4.2.2	Object Modelling .....	29
4.3	AHGF Maintenance Environment.....	30
4.4	AHGF Conceptual Modelling Methodology .....	31
4.4.1	Relevant Standards .....	31
4.4.2	Modular Data Model Development.....	32
4.4.3	Suggested Toolkit .....	34
<b>5.</b>	<b>AHGF Data Products .....</b>	<b>35</b>
5.1	Data Product Requirements .....	35
5.2	Data Product Specifications .....	37
<b>6.</b>	<b>AHGF Data Sets .....</b>	<b>39</b>
6.1	Surface Water – Hydrology .....	39
6.1.1	Characteristics .....	39
6.1.2	Node/Link Network.....	40
6.1.3	Example Surface Water Hydrology .....	41
6.2	Surface Water – Measurement Sites .....	42
6.3	Ground water – Measurement sites .....	43
6.4	Ground Water – Hydrology .....	45
6.5	Terrain .....	45
6.6	Key Catchments .....	46
<b>7.</b>	<b>Next Steps .....</b>	<b>48</b>
	<b>References.....</b>	<b>49</b>

## List of Figures

Figure 1 NHD Schema	7
Figure 2 NHDPlus Schema	9
Figure 3 GWML model for aquifers.	13
Figure 4 Conceptual modelling lifecycle	23
Figure 5 An illustrative example of a conceptual model and a derived implementation model	27
Figure 6 Geometry based modelling.	28
Figure 7 Object based modelling.	29
Figure 8 Packaging of core ANZLIC Harmonised Data Model components	33
Figure 9 Distinction between supply and product delivery in the AHGF	36
Figure 10 Example model of surface water - hydrology.	41
Figure 11 Waterbody features package.	42
Figure 12 An example of how monitoring points are represented within the AHGF conceptual model.	43
Figure 13 Example bore holes (Djokic 2008).	44
Figure 14 Bore hole model.	44
Figure 15 Demarcating drainage areas (Petheram 2008).	46
Figure 16 Drainage areas in the AHGF conceptual model.	47

## List of Tables

Table 1 AHGF Development Activities	19
-------------------------------------	----



## 1. INTRODUCTION

In order to manage a resource, it is necessary to have information about the features of that resource. For example, to manage a road system, information is required about the various roads within the road network and how these roads connect. Similarly, a land administration system cannot function without knowledge of the individual land parcels comprising the land system.

For water, the same applies – in order to manage water it is necessary to have knowledge of the features within the Australian water system. That is, the catchments, streams, aquifers, floodplains, storages, and wetlands that make up the system. Furthermore, to develop and deliver water resources information, location of these features is vital, as is information about the interactions between features.

The following document describes, in detail, the Australian Hydrological Geospatial Fabric (AHGF) with respect to the key components and methodologies required to build, manage and govern it. Development of the AHGF is one of the first attempts in Australia to build a suite of national data sets related specifically to water and with a requirement to support more than cartographic use cases. It is also one of the first attempts to build such a product capable of supporting evolution to the extent and in the time-frame required of the water domain.

### 1.1 Background

In 2007, as part of the Australian Government's National Plan for Water Security, the Bureau of Meteorology's (BOM) role was expanded to include management of Australia's water information. This new role requires the BOM to collect, manage and publish information from a wide variety of sources about a wide variety of aspects of Australia's water resources. The BOM will also be required to produce annual water accounts and water resource assessments from late 2009.

As part of this role, the BOM has recognised the importance of consistent framework data sets to support their new functions. Unfortunately, there are currently no nationally consistent hydrological data sets for Australia capable of supporting the BOM's requirements. That is, there is no consistent national stream network, the national digital elevation model (DEM) is of limited value for hydrological forecasting and accounting and data sets maintained by State and Territory agencies cannot be simply joined to produce a national picture. Those national data sets that do exist have been developed primarily for supporting cartography programs and hence little effort has been invested to ensure basic functions such as connectivity or consistent identification of features.

The initial implementation of the AHGF will focus on meeting the requirements of the Australian Water Resources Information System (AWRIS). This system is currently being developed by the BOM in order to support production of the various outputs required of it.

In time, the AHGF will become the foundation geospatial data set for the management of water throughout the nation, used by many public and private sector organisations for a range of purposes.

## 1.2 The Australian Hydrological Geospatial Fabric

The Australian Hydrological Geospatial Fabric (AHGF) is a suite of well maintained, evolving, authoritative data products containing a consistent representation of every feature, and the connectivity between features, of the Australian water system. These products are capable of supporting a wide variety of uses for many users. As such, the AHGF will become the framework geospatial information upon which Australia's water information related activities are based and through which they are related.

The concept of the AHGF relates to a baseline suite of common data sets that allow shared understanding of definition, naming and characterisation of particular landscape elements. For the initial implementation of the AHGF these baseline data sets should include: Surface Water - Hydrology, Surface Water – Measurement Sites, Groundwater – Measurement Sites, Groundwater – Hydrology, 1" Digital Elevation Model (DEM) and Key Catchments.

The definition of the AHGF above implies a number of broad architectural concerns regarding how the AHGF is delivered, to allow it to be realised from both a data maintenance perspective and to maximise its value to stakeholders. In particular, stakeholders will need to be able to predictably use the AHGF, and use it consistently over time. The AHGF itself needs to evolve as more aspects of the landscape are brought into consideration, more data products are delivered using it, or more sophisticated whole-of-landscape perspectives are implemented. It is not, therefore, adequate to create a one-off data model and rely on the user to interpret the semantics and business rules behind the content. Even a well documented data set imposes a considerable challenge if it is to be integrated or understood in terms of many other related data sets.

The AHGF itself is expected to extend from a core focus on enablers for water resource accounting to a richer, interconnected set of data that allows the relationships between landscape, hydrology and climate to be better understood, and improved prediction and management of water resource issues undertaken.

There is a requirement to understand how the AHGF can be deployed as a stable baseline, whilst undergoing evolution and maintenance processes. The key is effective modularity of the design, allowing for stability of components, whilst supporting update cycles where required. This is critical for the concept of the AHGF as an agreed baseline for multiple related activities to use.

Whilst theory and sufficient evidence of emerging best practice abound for modular software code development, the same is not true for data management practices. The methodology described here seeks to apply the lessons learned from data management, software management and commonalities in data modelling approaches to synthesise a practical methodology for AHGF design.

### 1.3 Use of the AHGF

The initial purpose of the AHGF is meeting the needs of the BOM and water information providers in fulfilling the requirements of the Commonwealth Water Act 2007 (Water Act) and associated Water Regulations 2008.

Details of specific uses and users of the AHGF can be described as a set of use cases, and can be helpful in determining requirements for the design of the AHGF. Key use cases for the AHGF include:

- Support generation of water accounts

The BOM is required, under the Water Act, to compile and maintain water accounts for Australia, including the National Water Account. The AHGF will be used to provide the contextual data (such as catchment boundaries) used to develop and produce these water accounts. It is expected that AHGF data sets will also be used by State governments in preparation of State and regional water accounts.

- Support generation of water assessments

As with water accounts, the BOM is required to provide regular reports on the status of Australia's water resources and patterns of usage as well as forecasts of the future availability of those resources. Nationally consistent geospatial information will be required to produce these assessments and forecasts, and to ensure that assessments can be compared over time.

- Ensure consistency in data supplied by water information providers

The Water Regulations 2008 identifies a number of individuals and entities that are required to supply water information to the BOM. In order for the supplied information to be of use in developing products at a national scale, it is vital that data be provided against a consistent and coherent baseline. The AHGF will provide this baseline framework ensuring data from different regions or organisations can be integrated effectively.

- Assist in dissemination of water information

The BOM's new role also includes responsibility for dissemination of Australia's water information. As with collection of information, the AHGF will provide a coherent framework upon which data sets can be published.

- Building comparable data sets

Ensuring data products are comparable across regions or time is important in developing long-term understanding of changes in water resource availability, use and management and in identifying trends. By utilising a common underlying framework of core data sets (for example a consistently identified national set of water features), the ability to compare and contrast results, forecasts and assessments will be enhanced. The AHGF will provide these framework data sets.

It is expected that as the AHGF is implemented, it will be used by a wider range of individuals and organisations for a variety of other purposes such as model development and calibration, planning and operational management of water resources. The expected diversity of users and uses requires that the AHGF be designed to accommodate multiple, potentially unidentified use cases.

### 1.4 Project History

This report is the culmination of a project to design the AHGF. The original deliverables of this project were:

- Conduct workshops with stakeholders and domain experts. Summarise and report on key outcomes from these workshops;
- A document specifying the design and function of Version 1.0 of the AHGF incorporating key findings from workshops conducted in the course of the project; and
- A technical report describing AHGF governance. This report will provide guidelines on the governance required to manage the evolution of the AHGF (Parashar 2007).

As a result of significant changes in the AHGF project landscape and maturation in understanding of the AHGF over the course of the project, the planned path to developing this report was not taken and some of the deliverables have changed.

In particular, the workshop elements, both stakeholder engagement and expert review, of this project have not been undertaken. This was due primarily to time constraints caused by delays at the project outset and the need to change methodology along the way. It is the opinion of the authors that this lack of undertaking has not affected the result to date. However, they are still necessary and should be undertaken in the near future.

A significant delay in the project was caused by the need to change project methodology part way through. It had been originally planned to test some of the methods describe herein. This necessitated the need to capture a set of test use cases. This capture process proved unsuccessful due to the inability to identify a suitable study audience within the project team.

### 1.5 Report Structure

This report describes the key components and proposes a method for developing the AHGF. Section 2 reviews similar data sets from the US, Canada and Europe. Section 3 describes a set of principles and proposes a methodology for developing the AHGF. Sections 4 and 5 describe two key components of the AHGF, namely the conceptual model (Section 4) and the Data Product Specification (DPS) (Section 5). Section 6 discusses the core data sets that the initial implementation of the AHGF should contain. Finally, Section 7 recommends next steps in development of the AHGF.

## 2. REVIEW OF HYDROLOGICAL GEOSPATIAL FABRICS

The AHGF is a collection of data sets representing all features in the Australian water system and the connections between those features. Its purpose is to provide reference data sets for data of hydrological significance. Reference data sets for hydrological data have been compiled by various organisations, most significantly in the United States and Europe. An overview of the key initiatives developing national or continental scale reference data sets for hydrological data is useful in determining appropriate methodologies for building the AHGF. This section provides such an overview, derived from two more detailed reviews by Lefort (2008) and Power (2007).

### 2.1 US Hydrological Reference Data Sets

Multiple government agencies in the United States collaborate to acquire, share and exploit water data (US GAO 2003, OSTP 2007, Cripe and Grabs 2007, Maidment and Tarboton 2007).

- The United States Geological Survey (USGS) manages the stream gauges for surface water monitoring (the National Water Information System – NWIS), the groundwater sensing networks (GWIS) and aims at real-time monitoring of water quantity (see <http://waterdata.usgs.gov/nwis/rt>) and continental scale aggregation of water status, including the USGS Climate Response Network for groundwater status (see <http://groundwaterwatch.usgs.gov/>), and the National Water-Quality Assessment Program (NAWQA) and SPATIally Referenced Regressions On Watershed Attributes Model (SPARROW) for quality trends (USGS 2008a, USGS 2008b).
- The Environmental Protection Agency (EPA) manages the environmental risks (pollution, river health) with data acquisition and processing systems focusing on water quality, with an emphasis on Total Maximum Daily Loads (TMDL) and impaired water zones (EPA 2008).
- A third major agency, the National Oceanic and Atmospheric Administration (NOAA) manages weather information, and leads the implementation (NIDIS Team 2007) of the National Integrated Drought Information System, a national drought information portal, which was launched in November 2007.

All these agencies have worked on the aggregation, transformation and dissemination of data. With an emphasis on water scarcity issues for USGS and more recently NOAA as well as on water quality issues (EPA). A common, national reference data set underpins many of the monitoring and assessment projects mentioned above. The USGS and EPA both primarily use the National Hydrography Data set (NHD) and its newer version NHDPlus as the source for geospatial reference data on surface water. NHD and NHDPlus are described in more detail below.

The reference data set for drainage basins / stream delineation traditionally used by NOAA is the Integrated Hydrologic Automated Basin Boundary System (IHABBS), but a migration to

NHDPlus is expected with the replacement of lumped models with distributed ones (IHABBS 2005, Smith et al. 2004).

### 2.1.1 National Hydrography Data set (NHD)

The National Hydrography Data Set (NHD) was developed as a cooperative effort by the USGS and EPA along with other federal, state or local agencies (NHD 2007). It is maintained by a number of state agencies under the overall stewardship of the USGS. NHD represents surface water information, natural and constructed, in a drainage network. The water feature information includes a classification (e.g. river/stream, lake/pond, canal/ditch), other characteristics (e.g. an indication if it is a perennial, intermittent or ephemeral water source), its geometry and related measures, geographic name, direction of water flow, and an assigned identifier or ‘reach code’.

NHD was based on the integration of the USGS Digital Line Graph (DLG) and Environmental Protection Agency’s Reach File. A water feature in NHD is represented as a point, line or area, and features are topologically connected where they intersect or overlap. Connectivity within the stream network is maintained through water bodies by including centrelines connecting streams into and out of (for example) lakes. Reaches correspond to water features within NHD. A unique reach code is assigned to each reach and remains permanently associated with that reach. If the reach is deleted, the reach code is retired and not re-used. Reach codes have been designed in such a way as to accommodate future inclusion of higher resolution data.

The data model for NHD corresponds with the data set and shown in

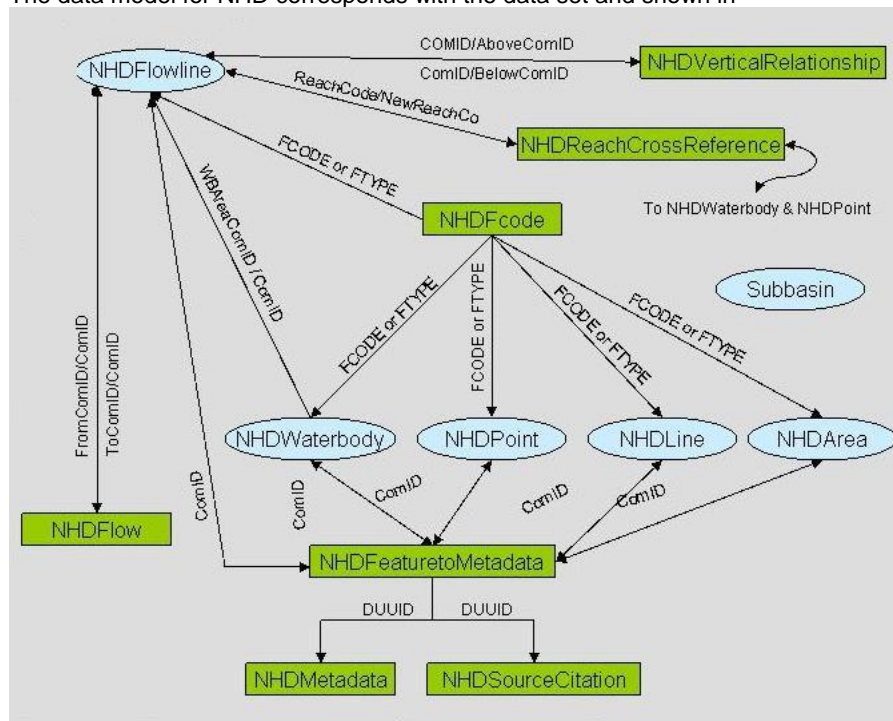


Figure 1 is a schema diagram depicting the components of the NHD data and their relationships. NHD can be delivered as a collection of files. The blue ellipses (water features) and green rectangles (associated features and relationships between features) relate to the accompanying



NHDPlus was built as a sequence of seven processing steps (NHD 2006b) summarised below:

1. Improve the NHD 1:100K Linear Network and Feature Attributes – the NHD surface water network was corrected where necessary by identifying errors starting at the sub-basin level and then combining and repeating for hydrologic regions and higher levels.
2. Computation of Value Added Attributes – 20 additional attributes are computed from the NHD by navigating the network both downstream and upstream.
3. Development of Catchments, Max/Min Elevations, Flow Accumulation, and Flow Direction Grids – the ARC/INFO GRID module ‘watershed’ was used with the NED Digital Elevation Model (DEM) to create catchments for each NHD flowline.
4. Development of International Catchment Areas – catchments that overlap Canada or Mexico require data sets not in NHD to more accurately calculate streamflow or other hydrologic attributes. The streams that receive flows from outside the United States were identified and extra data sets used to derive these values.
5. Development of Catchment Attributes – catchment attributes were computed using mean annual precipitation data, mean annual temperature data, and the National Land Cover Data Set (NLCD).
6. Development of Flow Volume and Velocity Estimates – flow volume and velocity estimates were calculated for each flowline using climate data and the NLCD.
7. Elevation Post-Processing, Smoothing, and Slope Calculations – the flowline elevations derived in step 3 were used to calculate the slope for all flowlines using various smoothing techniques.

After this process, the data was packaged and a final quality assurance/quality control (QA/QC) performed.

The NHDPlus data set is distributed as ESRI shapefiles (catchment geometry, hydrography, stream gauge) by hydrologic region and as ESRI Grids (flow accumulation and direction, catchment grid, and elevation) by ‘production units’, a sub part of the hydrologic region. Data versioning is indicated using a filename convention similar to the NHD. Supporting tools are provided to navigate the drainage network, upstream and downstream.

The NHDPlus data model and its relationship to NHD are shown in Figure 2.



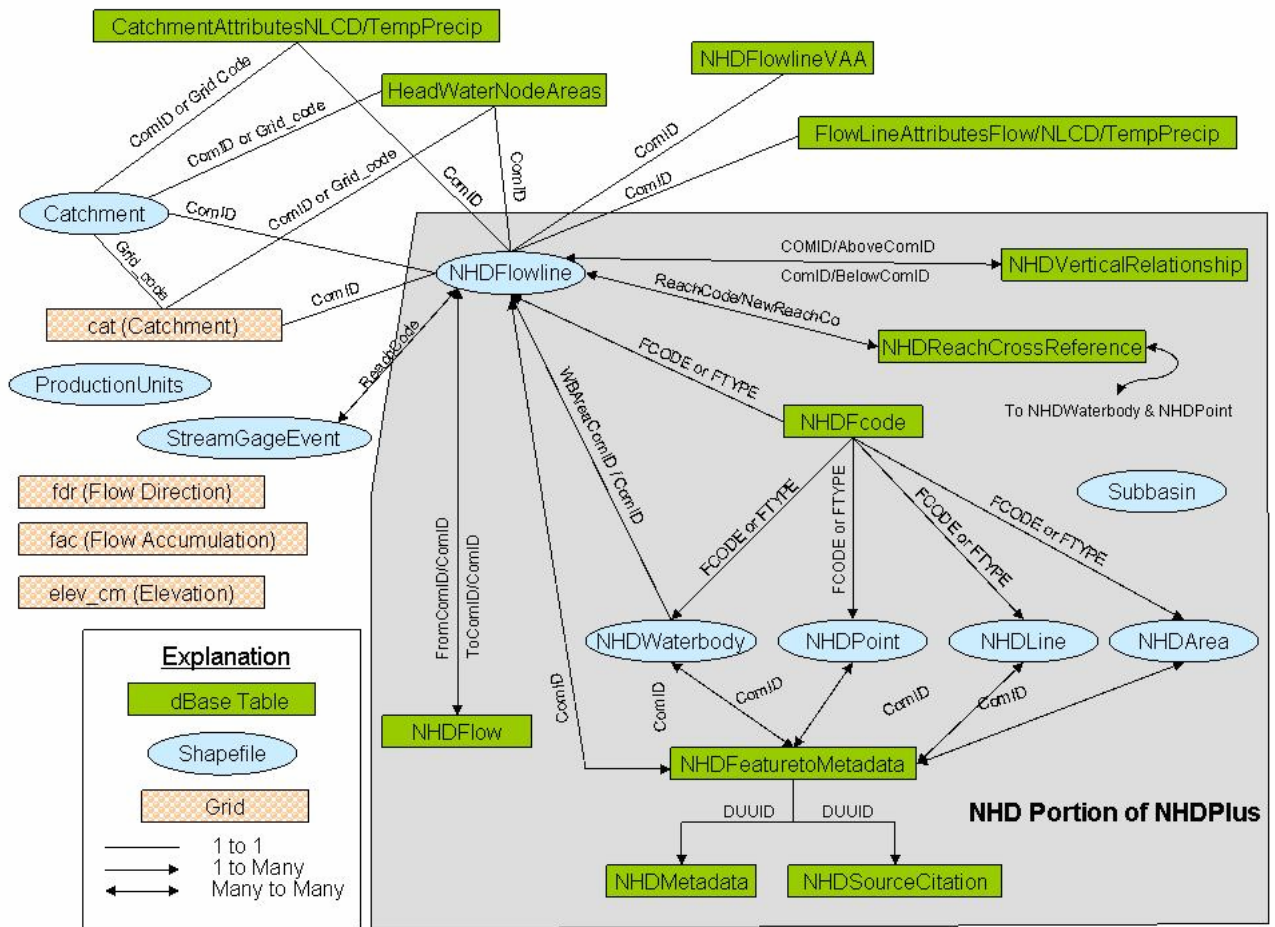


Figure 2 NHDPlus Schema

## 2.2 European Activities

The Water Framework Directive (WFD) is the key driver for projects in water information management in Europe. Implementation of the WFD is led by the European Environment Agency (EEA) and facilitated through various research projects. Several projects relate to reference data sets as described below.

### 2.2.1 CCM 2

The EU Joint Research Centre has assembled a reference data set for hydrologic features, the Catchment Characterization and Modelling data set (CCM), which is comparable to NHDPlus in the United States. The data set consists of a pan-European database of river networks and catchments. A second version of the data set, [CCM2](#), was released in July 2007 (Vogt 2007a, Vogt 2007b) and covers the entire pan-European continent, including the Atlantic islands, Iceland and Turkey. It includes a hierarchical set of river segments and catchments based on the Strahler order, a lake layer and structured hydrological feature codes based on the Pfafstetter system. The CCM2 data set should support the deployment of the Water Information System for

Europe ([WISE](#)) portal, a Europe-wide water data portal to give access to all the data collected for the different articles of the Water Framework Directive (WFD).

## 2.2.2 SANDRE

French initiatives in water information systems have led to the creation of a National Data Reference Centre for Water (SANDRE) which includes among its missions the dissemination of water-related reference frames and metadata. The three main organisations contributing to SANDRE at the technical level are:

- OIEAU: The Office International de l'Eau (IOWater) based in Limoges (OIEAU 2008);
- IGN: The Institut Géographique National (IGN 2008); and
- BRGM **Error! Reference source not found.**: Bureau de recherches géologiques et minières.

OIEAU and IGN jointly manage BD Carthage (BD Carthage n.d.), the reference data set for surface water. The reference data set for groundwater, BDRHF (BDRHF n.d.) is a joint project of BRGM and OIEAU. BDRHF applies an identification system for groundwater objects based on a rich conceptual model which has been evaluated against WFD requirements (Normand & Mardehl 2004) and proposed as the basis for the delineation of groundwater objects to other European parties involved in the BRIDGE project (BRIDGE 2005, BRIDGE 2006).

## 2.3 Other Initiatives

### 2.3.1 Arc Hydro

The Arc Hydro data model (Maidment 2002) and Arc Hydro tools (Arc 2001) were developed by the GIS Water Resources Consortium, a collaboration between the University of Texas at Austin's Centre for Research in Water Resources (CRWR) and the Environmental Systems Research Institute, Inc. (ESRI). The Arc Hydro data model describes natural surface water and associated information such as time series data for water quality and flow. The Arc Hydro tools are a collection of utilities that leverage the Arc Hydro data model and operate in the ESRI ArcGIS platform providing functions such as flow path tracing, watershed processing and terrain processing of a DEM. The tools are used to establish the required relationships and attributes in the GIS data structures, for example establishing the network topology and assigning measure attributes for a stream. The tools also provide the common base functions for water resource applications, such as watershed delineation.

The Arc Hydro data model categorises surface water resources information into five packages: Network, Drainage, River Channel, Hydrography, and Time Series. A subset of the Arc Hydro data model allowing capture of the most important geospatial information relating to water resources is also available as the 'Arc Hydro Framework'. While Arc Hydro concerns surface water only, a groundwater data model template is under revision and available from ESRI.

The five packages that form the Arc Hydro model are summarised below:

- Network – captures the connectivity and flow of water through streams, rivers and other water bodies;
- Drainage – representation of the surface water drainage system usually derived from land surface topography;
- River Channel – the shape of a stream or river using a series of cross sections that collectively represent the three dimensional view.
- Hydrography – map information for points, lines and other water body features.
- Time series – representation of the time varying properties of water features.

Due to development in close alignment with ESRI ArcGIS, the Arc Hydro data model has a well established implementation schema and can be extended with additional classes and/or attributes.

### **2.3.2 Arc Hydro Groundwater Data Model**

The ESRI Groundwater data model, an evolution of the Arc Hydro surface water data model, is a geodatabase design for representing multidimensional groundwater data.

The data model aims to achieve four goals (Strassberg 2005):

- Representing regional groundwater systems – a regional groundwater system is a large areal extent defining a volume below the surface through which groundwater flows;
- Representing site scale groundwater information – the purpose of a site scale representation is to define in detail a small region of interest within a larger system;
- Enabling integration of surface water and groundwater information – links between the surface water and groundwater are defined allowing water flow between these systems to be modelled; and
- Facilitate the connection of groundwater simulation models with GIS – classes are defined in the groundwater data model to represent the interaction with third party simulation models through the structure of the inputs and outputs.

While the Groundwater data model refers to the Arc Hydro data model to allow capture of surface to groundwater interactions, the two data models are, and will remain, separate. A toolkit to support the processing of Groundwater data sets, developed by Aquaveo in partnership with ESRI is due for release in late 2008.

### 2.3.3 Canadian Groundwater Data Model

The Earth Sciences Sector of Natural Resources Canada has numerous groundwater initiatives aimed at assessing groundwater quality and quantity (National Groundwater Database 2008). There are various projects underway to address these issues, one of which is to develop a National Groundwater Database (NGWD). This is a virtual repository of information in different data formats, all relating to groundwater in Canada. For example, there is information contained in documents, pictures, maps, well data, web sites, and web services. This information is managed using a web portal allowing users to perform simple keyword searches.

The NGWD is defined as a series of activities to:

- implement collaboration and develop a mechanism to collect groundwater information;
- design and implement data management architecture and standards to store and exchange groundwater information through internal diffusion channels and national initiatives and provide linkage with external and internal partners; and
- deliver information in a usable form to governments, educators, practitioners, and the general public, either as digital data or outreach and education material (National Groundwater Database 2008).

The NGWD does not manage a centralised database of information but rather maintains a catalogue of existing relevant data sources distributed across the country. The portal acts as a broker of information. User requests result in a list of matching sources in a similar manner as a web search engine. The NGWD is an example of a Service Oriented Architecture (SOA) where information repositories are accessed using a predefined protocol. Each data source supports interfaces allowing information to be searched and accessed.

A related activity is the Ground Water Markup Language (GWML) developed by the Geological Survey of Canada (GSC). GWML is an XML standard for the exchange of ground water data. GWML is based on Geoscience Markup Language (GeoSciML) and Observations & Measurements (O&M). GeoSciML is another markup language that was developed to support interoperability of geologic survey information such as geologic units, earth material, geologic structure, and bore holes. O&M is a data model for capturing observations and measurements made by machines or humans. These existing standards are extensively used in GWML since ground water exists as water in the ground and bore measurements are the main source of observation data.

GWML adds to GeoSciML and O&M features such as aquifers, water wells and attributes such as capacity. A UML diagram describing the aquifer portion of GWML is shown in Figure 3, from (NRC 2008).

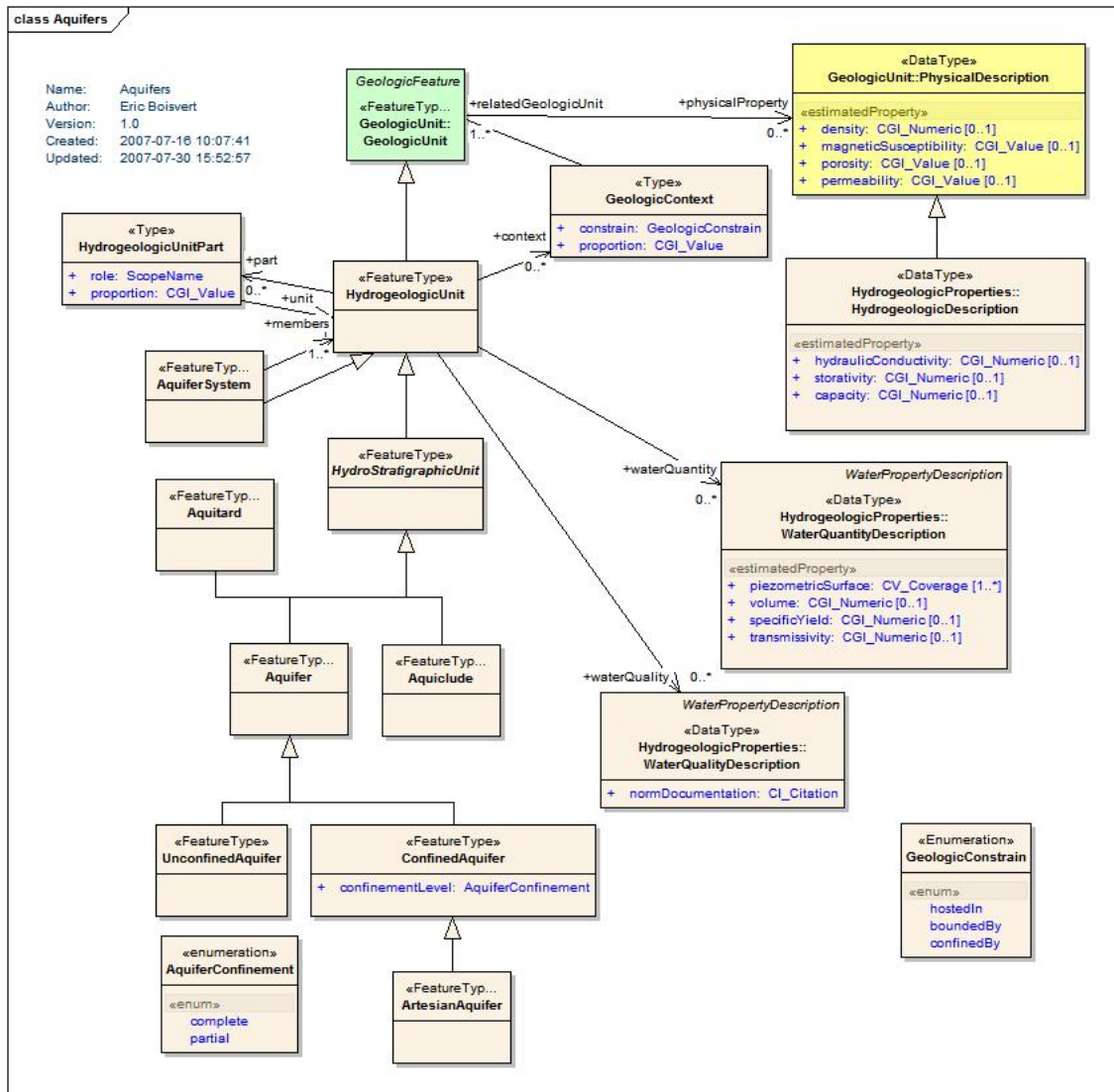


Figure 3 GWML model for aquifers.

## 2.4 Discussion

Despite the obvious differences in approach, the various data products and data models discussed in this section can all be regarded as attempts to build hydrological geospatial fabrics or parts thereof. For the AHGF, it is important to learn from each of these exercises to ensure that the AHGF is an improvement on recognised best practice rather than an implementation of the status quo.

One difficulty in evaluating the different data products and models is the lack of consistency in how the efforts are documented. Data products are documented at different layers of abstraction and different levels of detail making comparison between the various ventures problematic. This indicates a need for a consistent documentation methodology to be developed.

A key lesson to take from a number of these initiatives is the risk in not separating conceptual models from the requirements of particular implementations. For a number of the data products reviewed here, the information model has had a focus on implementation. For example, with NHD and NHDPlus, the data model is described in terms of implementation details such as shapefiles and database tables. The product is built by segmenting the landscape into identified pieces at a particular scale, with further detail added subsequently. The model lacks any higher level abstraction or conceptual representation. In the Australian context, this upfront delineation of catchment boundaries is problematic due to the nature of the terrain.

In a similar way ArcHydro, although having a conceptual model, is tied to a particular implementation by upfront selection of scale and representation of objects. One outcome of this is that features cannot be viewed in multiple ways, for example at various levels of detail or complexity.

This focus on implementation without a higher level of abstraction results in models that limit the flexibility of data products to support many use cases. That is, the model will only be efficient for certain use cases and will require an expensive transformation process in order to support other use cases.

In the case of the Canadian Groundwater project, the focus has primarily been on the development of a conceptual information model and a related XML encoding for data transfer. There appears to be little effort on the development of an authoritative suite of data products. For the AHGF, there is a great need for such products.

Another interesting point is the contrast between NHD and ArcHydro in terms of the way in which the system is designed to support maintenance and evolution. The design of ArcHydro supports exploitation of the system as designed. It has limited provision for evolution of the networks or ongoing maintenance of the data. In contrast NHD has a strong process model with considerable detail on processes for building, maintenance and extension. While the specific processes in NHD may not be directly applicable in the Australian landscape, the methodology and level of documentation detail may provide useful starting points in the development of the AHGF.

In each of the data products and models reviewed above, identity management is a key concern. The ability to uniquely identify features within the landscape is a common requirement, and management of an identification hierarchy that allows for data at multiple levels of detail to be maintained is featured in several of the systems including NHD and CCM2. NHD also considers the issue of identifier management over time and evolution of data products which will be important in ensuring the AHGF can support an evolvable picture of the landscape.

### 3. METHODOLOGY

Modelling and capturing data with respect to some aspect of the landscape, such as hydrology, is a large and expensive task. If this task is not undertaken in an organised fashion, the risk of failing to meet the business requirements that the data has been captured to support is high. The following section proposes a set of principles and a methodology for systematically designing, developing and implementing the AHGF.

#### 3.1 AHGF Development Principles

A significant body of literature exists regarding software design cycles (cf. [http://en.wikipedia.org/wiki/Software\\_development\\_process](http://en.wikipedia.org/wiki/Software_development_process)). The principles for developing the AHGF draw upon proven best practices, such as Design By Contract (DBC) and AGILE programming methodologies. That is:

- tackle the major risks first;
- iterate rapidly, with end user involvement;
- design interfaces between components (the contract); and
- develop unit tests so that each piece of the implementation can be tested, and key behaviours maintained across update cycles.

The methodology proposed for development of the AHGF introduces a number of further principles relevant to the specific issues facing data management and supply functions. These are:

- Identification of appropriate principles for disaggregation of the process into data management units, allowing a staged approach to data set development without invalidating links to the AHGF by data using intermediate products;
- Recognition that the problem has many aspects, and requirements analysis and best practice are incomplete, hence the need for an implementation cycle that presumes that issues will emerge and be resolved as, and when, necessary;
- Separate concerns of data maintenance activities from the need to support many derived data products;
- Separate the conceptual model from various implementation views; and
- Identify clear governance patterns for identifier management, update and distribution policies;

The remainder of this section discusses a number of these principles.

### 3.1.1 Risks

In the case of the AHGF, the key risks to be addressed will include:

- ensuring the interpretation of the real world during data capture is consistent with end user expectations;
- ensuring the level of detail of capture is appropriate for initial use;
- ensuring that evolution (to provide, for example, improved detail or a more complete model for a particular area of interest) does not invalidate the underpinning agreed common conceptual model or invalidate cross references in other client data sets;
- identifying those historical views of the data that need to be maintained and accessed and ensuring this is possible; and
- ensuring users can adequately formulate an access request, understand the results and integrate the process with their own toolkits.

Addressing each of these risks will minimise the overall cost and time to delivery of the AHGF and the AWRIS system over the medium and long term. Following the proposed methodology also allows short-term investment to be protected through early identification of aspects that require stability over the lifetime of the AHGF as an enduring asset.

### 3.1.2 Staged Approach to Data Set Development

A key principle for developing the AHGF will be to use a staged approach, with each stage enabling consistency in reporting on some particular aspects. For example, the first step in realising the AHGF should be the capture of a set of hydrologically coherent catchment areas for the nation. Detailed stream network information can then be captured for each of these in a rolling program. More complex tasks, such as modelling surface/groundwater interactions and detailing flows within a stream network can then follow as the required understanding and knowledge of these aspects becomes available.

To support a phased data capture and improvement cycle, the AHGF design must provide for:

- spatial disaggregation of the data set into units with identifiers that are likely to remain stable across improvements;
- varying levels of detail within the information model, to cope with both variability in available data at inception, and the potential need to develop a rolling program of improvements;
- early introduction of versioning and data lifecycle in a fashion consistent with the broader AWRIS requirements.



### 3.1.3 Iterative Development and Rapid Prototyping

A number of factors associated with the AHGF suggest that rapid prototyping and a comprehensive end-to-end validation process will be required. These factors include:

- the complexities of the domain, such as land/water and surface/groundwater interactions;
- the level of investment required to realise a consistent data set;
- the paucity of experience in delivering some data sets to client applications and environments; and
- the need to support evolution in terms of functionality, detail, accuracy and accessibility of a data set.

### 3.1.4 Separation of Conceptual and Implementation Models

The AHGF will need to support a wide variety of use cases and hence a wide range of applications. It is unlikely that a single data model can be developed that is capable of efficiently supporting these individual applications. As such, a key component of the AHGF will be a rich, modular and extensible conceptual information model, capable of describing known behaviours of hydrologic features, and becoming richer as more characteristics are determined and understood. From this relatively comprehensive model, simple, consistent and stable data products can be derived to support various use cases and application.

Progress towards such a model and its role are described in Section 0.

### 3.1.5 End-User Involvement

In most cases, the AHGF data sets will be evolved from existing data products. Even in cases where data sets are recaptured in detail, existing users of the data will have particular views that need to be accommodated in the generation process.

The end user has a need for stability, and will in general only be interested in certain aspects of the data. The complexity of versioning, QA/QC, licensing and other aspects of the data is unavoidable, but the AHGF design methodology needs to ensure that users are not faced with a plethora of similar, but potentially subtly different, approaches to these aspects. The AHGF needs to commoditise data access, within a consistent framework that can operate on trust rather than laborious examination of complex documentation.

## 3.2 Proposed Methodology

The purpose of the AHGF methodology is to guide the design of the AHGF components through an end-to-end analysis and prototyping of the entire data production, management and distribution lifecycle. Consideration is also given to how component design can be validated in the broader AHGF framework.

For an effective data management process it is necessary to ensure:

1. capture of the most complete, available data; and
2. data supports the delivery of useful, derived data products.

It is also necessary to drive validation of the AHGF with real data to ensure the framework supports the operations it is designed for.

The table below shows a set of proposed activities for the development of the AHGF. These activities are implied by the scope of the AHGF implementation requirement: a deployed, maintained data set that can be used to support spatial referencing and disaggregation of other data products in the AWRIS system. The activities have been broken down according to the nature of the task and skill sets required. There is, in general, an interdependency between each task and its neighbours. Each should be undertaken with a mind to the principles described in Section 3.1.

Table 1 AHGF Development Activities

		Common Architecture		
Identify Source Data	Supply Contract (DPS) + SLA)	AHGF Conceptual Model	Service Design	Delivery Product Definition (DPS)
		AHGF DB	Service Implementation	Client tools (e.g. data download to Arc Hydro)
		Population	Deployment	Evaluation
Identity Management				
Technical Integration				

Tasks are colour-coded as follows:

- Key outputs of AHGF design process.
- Implementation feasibility and optimisation.
- Validation of data integrity, processing feasibility and value.
- Validation of AHGF outputs and support for stakeholder evaluation.
- Critical tasks to ensure AHGF meets AWRIS needs.

These tasks will be iterated, with the need for refinements driven by lessons learned from data supply experiences, technical implementation feasibility and end-user needs.

It important to note that some tasks are considered to be cross-cutting:

- Technical Integration – developing testable integration strategies to ensure that components are consistent with each other. This needs to be undertaken at all phases – design (of tests), unit testing implementation, and end-to-end (integration) testing of the entire system.
- Identity Management – as noted in the review of similar frameworks (eg. NHD and NHDPlus), this task is a critical and specialised role, ensuring that all aspects of the system design and data can share common identifiers across the end-to-end operation of

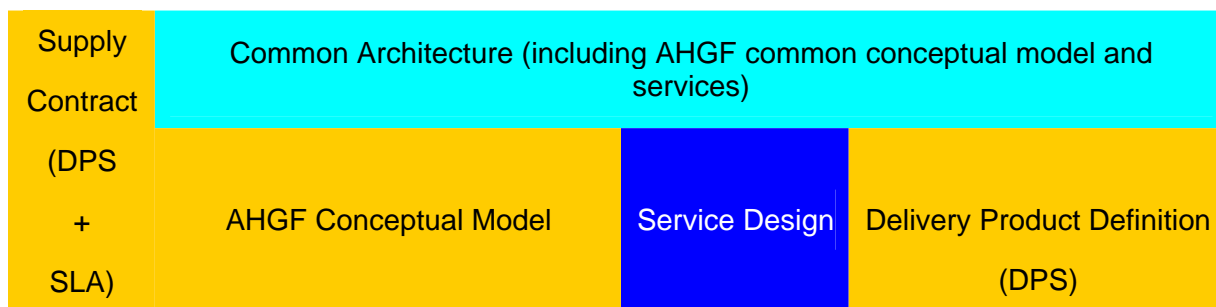
the system, including data maintenance and re-supply. It is expected that education of stakeholders on how object identity is managed in different cases may be required.

This methodology can be further broken down, horizontally into similar types of activities. These are: design, implementation and testing with real data. They can also be viewed vertically, to look at the complete set of tasks required to implement a particular component. The following sections describe each activity through the horizontal viewpoint, and then the vertical viewpoint.

### 3.2.1 Design phases

This layer of related activities identifies the design phases, from the perspective of what the AHGF can offer to users and applications.

This view separates out the distinct activities:



#### *Supply Contract*

This component is the key enabler of the AHGF. Without input data, the AHGF’s usefulness will be severely limited. This activity ensures that appropriate agreements are put in place to ensure data is available. Key artefacts of this activity are Data Product Specifications (DPS) (described in Section 5) and Service Level Agreements (SLA) between data providers and the owners of the AHGF.

#### *Common Architecture*

The common architecture provides an enabling activity that allows different aspects of the system design to share common definitions and agreements. Cross cutting issues such as versioning, identifier management and metadata reflect the architecture of the AHGF, and its role within AWRIS and other client systems.

The common architecture activity provides a home for all decisions which need to be shared by different aspects of the design. One approach to execution of this task is to derive the architecture by stratifying the AHGF product and packaging decisions required as either AHGF specific, or broader framework data set concepts. The common architecture is thus defined to be all aspects of the system that are not specific to a particular data product.

Appropriate documentation methodologies for the common architecture require further exploration.

### *AHGF Conceptual Model*

A key design principle of the AHGF is to separate the conceptual model from various implementation models. This allows for many implementation models, tuned to support particular use cases, to be developed. The AHGF Conceptual Model and its development are discussed in Section 0.

### *Delivery Product Definition*

The output of this activity is a set of data product specifications (DPS) that meet the requirements of the AHGF users. This activity involves use case and scenario analysis, in consultation with stakeholders.

Prototyping phases related to this activity are recommended to support the stakeholder engagement process.

More detail on data product specifications is provided in Section 5.

### *Service Design*

The design of services to access the AHGF will be driven by the design of data products and the use cases to be supported.

A key factor to be addressed in this activity is understanding the questions that can be asked by the client. These questions typically embed prior knowledge of the content and semantics of both the data and service interface, and are thus made possible through the common architecture component. Both user and provider (client and server) need to understand the system to an extent, and the service interface encapsulates that common understanding.

### 3.2.2 Implementation

The principle of iterative development and rapid prototyping is used to validate the design of individual components. With respect to the development of the AHGF, it will be necessary to have an early focus on development of the AHGF maintenance database, tools for exposing AHGF data and a small set of client tools.



#### *AHGF DB*

This task initially involves creation of database schema to support the data maintenance environment. Ideally this will be modular, with schemas for key packages in the reusable AHGF conceptual model.

It is expected that a data warehouse would also be maintained. This would be designed to support efficient and robust access via the queries required by the data product derivation process, in particular through online service interfaces.

#### *Service Implementation*

It will be necessary to test the feasibility and capability of the services required by the data delivery use cases. This step is iterative, with new technologies potentially becoming available to improve aspects of the implementation.

Support for multiple formats, transfer paradigms and security models may require multiple technologies to be deployed in parallel.

#### *Client tools*

The final validation of the AHGF will be its ability to deliver on its intent of providing a common baseline for references to the hydrology of Australia. In practical terms, this means ensuring data can be delivered to the environments where it will be used, preferably with the opportunity to be maintained in an ongoing fashion via automated updates.

One key target environment is expected to be Arc Hydro, enabling the application of available modelling tools to the set of agreed features maintained by the AHGF.

### 3.2.3 Data maintenance environment testing

This view of the matrix of activities focuses on the ability to create a database environment to manage an AHGF data product and its update processes. It is expected to be an iterative process, with challenges importing real world data cases leading to appraisal of the conceptual model, and refinements to the database implementation.

Task	Description	Output
AHGF Conceptual Model	Developed according to ISO conceptual modelling principles, with a focus on modularity and re-use.	UML model, with re-usable components clearly packaged.
AHGF DB	A physical database implementation, realising the model with tables, relationships and views (for example).	Database schema. Ideally multiple related schemas.
Population	Testing of the database design with real data.	Database.

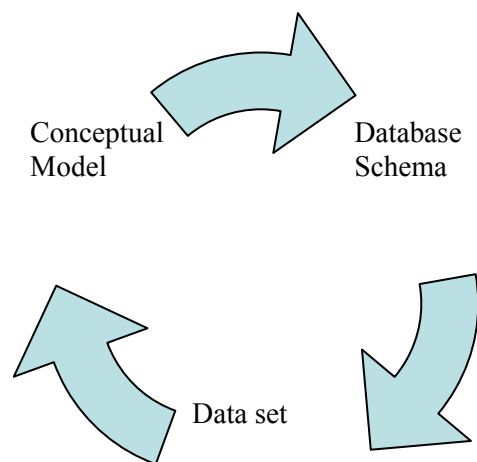


Figure 4 Conceptual modelling lifecycle

Issues with data not being able to be adequately expressed in an implementation model (for example within a database schema) should cause a review of the conceptual model to be triggered. Any redefinitions from this process should then flow to any derived data products. This cycle (shown in Figure 4) should be iterated often enough to have confidence that new applications will not require changes to ensure backwards compatibility.

### 3.2.4 Service interfaces

The creation of service interfaces to access the AHGF follows a similar iteration cycle to the conceptual model. However in this case, it is expected that service interfaces will impact mainly on the implementation of the database, in particular the need for de-normalised, read-only views designed to provide secure read-only access and efficient support for common access queries.

Task	Description	Output
Service Design	Specifications of the questions and responses that will be supported.	Formalised Service Profile documentation (see WRON-RM).
Service Implementation	Test, and if necessary refine software capable of implementing service interface.	Software reference implementation demonstrating feasibility and allowing further testing.
Deployment	Configuration of services against actual data.	Deployed test services for performance, functionality and integration testing.



## 4. DEVELOPING A CONCEPTUAL MODEL

The AHGF will be required to support a large number of diverse use cases. The BOM itself will use products from the AHGF to support: development and delivery of water accounts and water assessments, ensure consistency in information supplied by data providers and support delivery of information to others users. Other users will use the AHGF to support different use cases.

While it may be possible to support all of these use cases using a single implementation model, such as the Arc Hydro data model, it is unlikely that such a model will provide efficient support for all uses. Adoption of a single implementation model risks compromising operational efficiency of individual applications. For example, an application that supports ongoing management of a particular part of the water system may require a temporally focussed data model whereas an application for reporting at a national scale may require a more spatially oriented model.

A conceptual model provides a common framework that can support many different implementation models and allow for identification of relationships between representations of features in these models.

The following section discusses conceptual and implementation models and their roles in the AHGF. A number of options for developing a conceptual model are explored and finally a methodology for developing the AHGF Conceptual Model is proposed.

### 4.1 Conceptual and Implementation Models

A conceptual model is an abstraction of phenomena, or features, in the real world. It describes the essential features, captures their relationships, and may also represent behaviour. In contrast, an implementation model defines the structures, relationships, and behaviours needed to support an operational system.

Conceptual models allow for representation of features in different ways and are not constrained by implementation realities. Implementation models, on the other hand can be constrained by the functionality of the target system. For example, a conceptual model might allow for multiple geometric representations (at different scales, or using different methodologies) of a particular feature type. (e.g., a dam may be represented by a polygon (perimeter of water body) or a point location (the centroid)). Many GIS products, in which an implementation model would be used, require a single geometric representation for features. Users would select the geometry type that supported their particular use case.

Conceptual and implementation models are logically distinct from one another. There is, however, a mapping from the general conceptual model to a specific implementation model. In addition, as the implementation models are derived from a common conceptual model, coherence between implementation models is maintained. This coherence allows for consistent mapping between implementation models, ensuring semantic integrity, and a basis for structured extension and exploitation of the models. Instances of specific concepts that exist in derived implementation models maintain their unambiguous equality though use of a suitable

identity scheme. For example, a reach in a node link network is semantically identical to a reach in a blue line network if they are derived from the common conceptual model and reach ID1234 is identically the same reach in both realised implementation models.

The goal is to develop a single conceptual model that can be mapped to different implementation models, as required. These implementation models will be structured to specifically support a particular use case.

It should be noted that it is possible to directly implement a conceptual model as a lossless database schema, or XML transfer schema – i.e. one that contains all the elements in the conceptual model without redundancy. This can be considered the case for NHD, NHDPlus, and Arc Hydro where the conceptual model is derived from geodatabase structures (the implementation model). It is likely that the implementation of the data maintenance environment will tend to be closer to the conceptual model than any derived data product.

An example of the relationship between a conceptual and an associated implementation model is shown in Figure 5. A reach in the AHGF will have many complexities such as high-flow linkages and groundwater interactions. Figure 5 shows how a rich conceptual model provides for an extensible ability to capture different aspects of the real world phenomenon. On the other hand, implementations may have very specific capabilities and purposes, and be simplified, for example for cartographic display or spatial analysis of the proximity of other features at a specific (e.g. 1:250,000) map scale.



such as a database, when compared to standard numeric or character string data types. For these reasons, spatial data is often distinguished from attribute data and managed separately.

Since spatial data is considered special, spatial data modelling would begin by identifying the spatial features and categorising them in terms of their dimensionality: points, lines, polygons. Features with similar spatial characteristics would be grouped together as collections of like geographic features and organised as layers.

For example, a bore has a geographic point location and will have associated characteristics depending on its use, examples being water extraction, or environmental assessment. A bore layer would then be modelled as a feature data set consisting of a collection of points (the bore locations) and other common attributes (date installed, by who, method of construction, depth). Recorded observations from the bores are then managed as a time series associating measurements to the appropriate bore.

The focus on geography is embedded in the way the data is handled. Models are developed in terms of points, lines and polygons. Real world phenomena are categorised by their spatial description. In the NHD for example, a reservoir can be modelled as a point (as a landmark feature), or a polygon giving its areal extent. Other point features (e.g. gauging station, bore hole, piezometer,) are grouped together since they have the same spatial dimensionality. The linear and areal features are similarly organised.

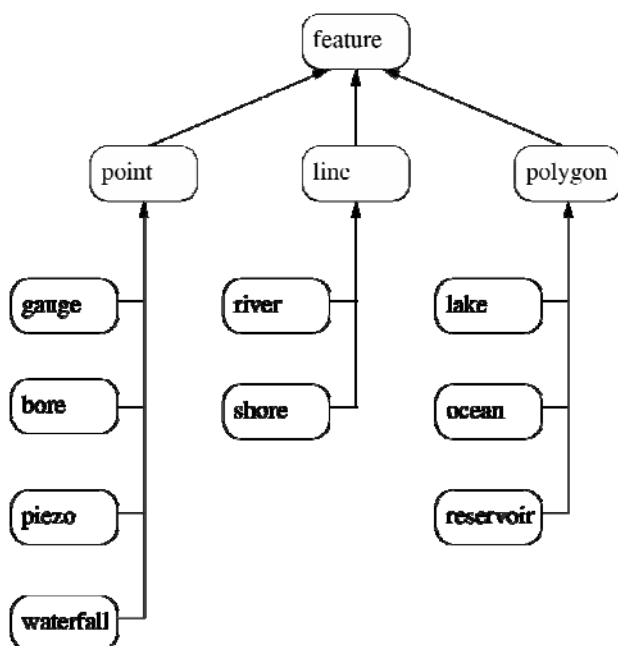


Figure 6 Geometry based modelling.

In the example shown in Figure 6 **Error! Reference source not found.**, the top level item, *feature*, corresponds to a real world phenomenon being modelled that has a geographic location, sometimes called a *geographic feature*. Then all features are categorised by their spatial dimensionality: points, lines and polygons. Those features that are points are grouped together.

The NHD and NHDPlus data sets model water features this way, categorising everything into a point, line, or area. A feature code is then used to define the feature type. The codes are references to a lookup table. This is a common modelling practice in the GIS community. One limitation of this method is that multiple views of a feature are not supported – the same lake cannot be represented as both a point feature and as a polygon.

### 4.2.2 Object Modelling

The alternative approach is to model real world phenomena as *objects* without special attention given to spatial characteristics. This is the method used in object oriented design. Figure 7 provides an example, an alternative to Figure 6. Here the features are categorised by their function, not their spatial characteristics.

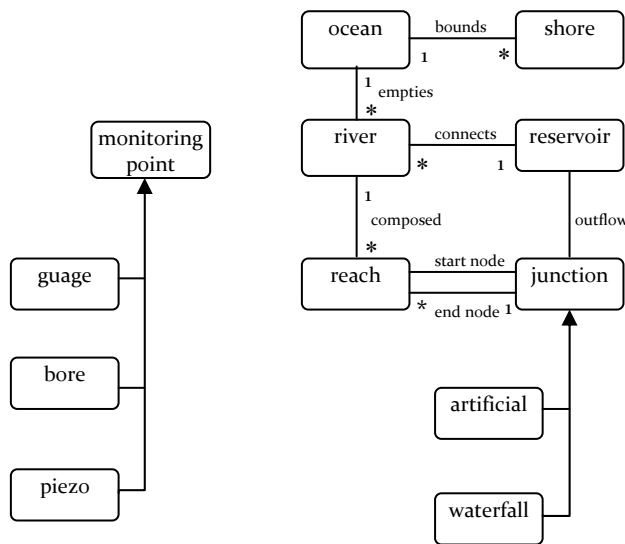


Figure 7 Object based modelling.

The obvious difference between Figure 6 and Figure 7 is that there are no *point*, *line*, or *polygon* features in Figure 7. The model is now focused on the things being modelled, and is not concerned with their geometric representation. The spatial description of a feature, for example a reservoir, is an attribute of the feature, not a first class concept. That is, there is no point/line/polygon box in the diagram under which the reservoir must reside. This means the model can accommodate a reservoir as having any kind of geometry, or even multiple geometries (e.g. a centroid, a high water mark, a low water mark). This is in contrast to Figure 6 where a reservoir must be a polygon.

Looking at some of the other aspects of Figure 7, a *gauge*, *bore*, and *piezo* are examples of a *monitoring point*: each records environmental information about a location. A *river* is composed of one or more *reaches*. A reach has a start and end node, defined as a *junction*. There can also

be two further kinds of junctions: *artificial* or *waterfall*. A *reservoir* and *ocean* both have relationships to a *river*, while a *lake* is shown to be separate.

Note: Figure 6 and Figure 7 are informal diagrams depicting the relationships between features. The arrows loosely correspond to inheritance or an ‘is-a-kind-of’ relationship. For example, in Figure 6, a *bore* is a kind of *monitoring point* which is also a *feature*. Attributes that are common to all ‘below’ items are included in the ‘higher’ item. Every feature may have a unique identifier and so this would appear in the *feature* item. Distinguishing attributes appear in sub-items.

The lines without arrows show named associations which sometimes include a cardinality (also called the multiplicity), indicating the number of items in the association. For example, a *river* is composed of many *reaches*. The association name applies both ways, so a *reach* is part of one *river*. The default cardinality is 1-1. For example a *reservoir* has one outflow *junction*, although this is not always the case in the real world, this is what the example model states. Note that a *reach* can only have one start node *junction*, but the same *junction* may be an end node for many *reaches*.

Diagrams such as these are used during initial brainstorming of ideas for developing models. They are then revised and formalised into UML.

Figure 7 is a simplified example of an object model for these feature types, and is not definitive. Alternative models could be arrived at using the same approach (for example, a monitoring point could be considered a kind of junction; a reservoir a special kind of lake), and more relationships defined between features (connections from a river/reach to a lake).

### 4.3 AHGF Maintenance Environment

It is expected that the AHGF will be established as a collection of water data sets and related information managed in some form of geodatabase. A geodatabase is the term given to a Relational Database Management System (RDBMS) that can operate on geographic data as efficiently as traditional numeric and character string data types. This requires spatial data support for its life cycle within the RDBMS: loading, indexing, querying, and maintenance.

Creation of the AHGF maintenance geodatabase will use the conceptual model in conjunction with the capabilities of the chosen RDBMS to derive an appropriate implementation model. This process may result in an implementation model that resembles the geometry based model of Figure 6 rather than the object based model of Figure 7.

For example, if the AHGF was to be created as an instance of the ESRI geodatabase then the Arc Hydro data model might be adopted as the starting point for the implementation model. Arc Hydro partitions features into points, lines, and polygons, and so there would need to be a process of aligning the features in the conceptual model to be compatible with the Arc Hydro structures.

The value of the conceptual model is that it is not influenced by the limitations of a chosen geodatabase implementation. Model development can focus on the phenomena being examined, how they are to be used, represented, and their relationships to other phenomena.

## 4.4 AHGF Conceptual Modelling Methodology

### 4.4.1 Relevant Standards

The methodology used for documenting the AHGF design should be based on the ISO 19000 series of standards, which represent best practices from the international sphere. Application of ISO standards is further guided by the implementing guidelines methodology (INSPIRE 2007)] developed by the European Commission's INSPIRE programme, which aims to apply best practices from ISO and elsewhere in the context of environmental information in the EU.

With respect to the AHGF Conceptual Model, ISO 19109 describes a General Feature Model and provides patterns and rules for allowing such processes as deriving implementation schemas from the conceptual model (e.g. UML to XML).

The conceptual model should be expressed in Unified Modelling Language (UML) diagrams, using a formal idiom that makes clear what relationships are intended to express.

UML as a modelling framework has distinct advantages over entity-relationship (ER) database schemas including:

- Object definitions can be derived from common baselines (specialisation);
- Inter-object relationship semantics can be made explicit, and described in detail (ER only allows for foreign key constraints);
- UML is naturally package or module oriented. In contrast, database schemas do not easily allow a standard schema to be specialised – schemas are self-contained implementations;
- UML supports modelling at several meta-levels (c.f. ISO 19763 Framework for Meta-model interoperability);
- UML is technology independent, whereas ER schemas are typically bound to a technology and data types it supports; and
- A UML model can be implemented in ER, XML schema, OWL ontology or other formats.

#### **4.4.2 Modular Data Model Development**

An important principle to use when developing the AHGF Conceptual Model is to employ a modular approach. That is, to concentrate only on those components that are specifically relevant to the domain.

To illustrate this, the ANZLIC Harmonised Data Model (HDM) is used as an example. Note that HDM deliberately avoided creating a module for hydrology, in anticipation of an activity such as the AHGF providing a well designed and supported national standard.

Figure 8 shows how the ANZLIC Harmonised Data Model builds common packages from relevant ISO standards. Note that basic data types are also defined in ISO packages, but these are realised by mapping to implementation data types, rather than importing data model elements.



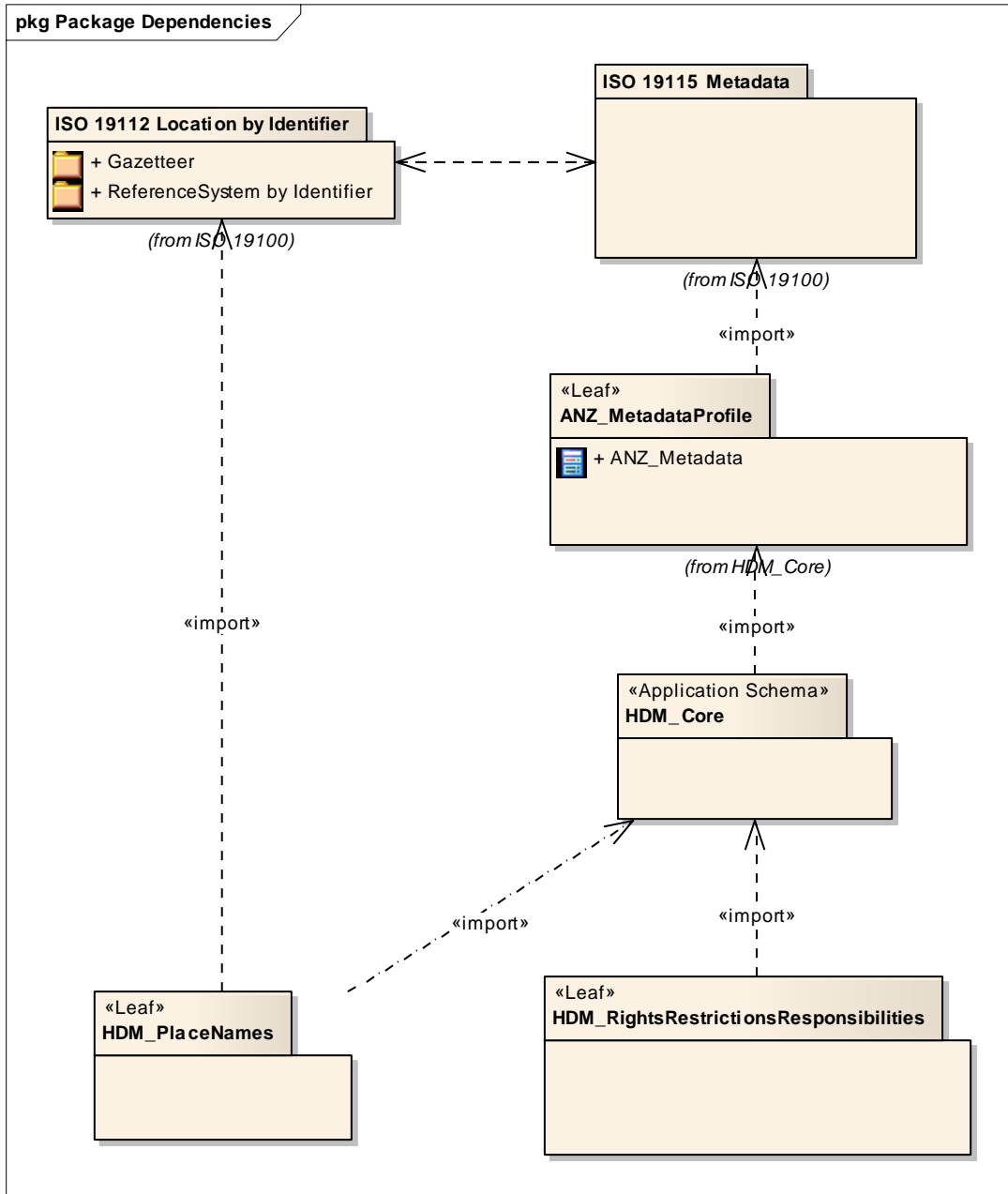


Figure 8 Packaging of core ANZLIC Harmonised Data Model components

In the water resources domain, the AHGF Hydrology module (sharing common versioning and metadata patterns with other AHGF data sets) might be referenced by a WaterRights package, which could share common metadata about the lifecycle of allocation with other types of rights through the HDM RightsRestrictionsResponsibilities module.

This would support, without necessarily dictating, linkages of water rights to landscape and cadastral elements, and one could ask a question “what rights are associated with this piece of the landscape?” and get water rights, carbon offsets, easements, planning zones and other information using a single tool. This may not be an immediate priority, but illustrates the

conceptual basis for achieving integration between the AHGF and its client applications through the ability to import the AHGF conceptual model in a standard way.

### 4.4.3 Suggested Toolkit

To facilitate a common approach to data modelling, across local and international activities, modelling should be undertaken using ISO and INSPIRE adopted practices. Based on the local experience of developing the ANZLIC Harmonised Data Model and international Geosciences Markup Language (GeoSciML) activities, it is highly recommended that the Hollow World toolkit for ISO data modelling is applied (SEE GRID 2007).

Within this toolkit, standardised data types are drawn from ISO standards. Automation of GML-schema representation (to provide a standards-based XML transfer format) is available via two different toolsets (ShapeChange – see <http://www.interactive-instruments.de/ugas/> and FullMoon – see <https://www.seegrid.csiro.au/twiki/bin/view/AppSchemas/FullMoon>). A key benefit of the HollowWorld approach is in-built support for key elements of the ISO UML profile. Another is the availability of version control mechanisms for packages.

## 5. AHGF DATA PRODUCTS

At the core of the AHGF is data. That is, data in the form of inputs to the AHGF and data in the form of products to be delivered to one or more of the many users and use cases that the AHGF must support.

Users of the AHGF must have confidence in the data. This applies, not only to users of AHGF outputs, but also to the owners of the AHGF, who must have confidence regarding input data sets.

The following section discusses two key components of the AHGF that are required in order to ensure users can have the required confidence in the various AHGF data products. These are the data input and delivery mechanism and data product specifications.

### 5.1 Data Product Requirements

The AHGF Conceptual Model will need to express complex interactions between features, and support the need to produce data at varying levels of detail. Individual data products, on the other hand, should be derived on a consistent basis, to meet the needs of end-users, who, in general, will only be interested in certain aspects of the entire model.

In order to successfully implement such a design, the realities of data management and processes involved in publishing data products must be considered. Best practice in conventional information management systems is to maintain a separation between transactional data management processes, and stable views of the data. Further discussion at: [http://en.wikipedia.org/wiki/Database\\_management\\_system#Logical\\_view\\_and\\_physical\\_view](http://en.wikipedia.org/wiki/Database_management_system#Logical_view_and_physical_view).

The AHGF must provide nationally consistent data products fit for specific purposes. For example, there will be a need for products that support water accounting and products that support water resource assessments. Such data sets will provide a view of the AHGF at certain map scales, which do not change if some parts were also available at more detailed scales. Likewise, it is unlikely that such products would contain all the versioning information required to maintain reproducibility of data products since end-users desire a simplified logical view, whereas the data management environment needs the ability to capture and manage all relevant information.

Figure 9 shows a conceptualisation of a proposed delivery architecture for the surface water component of the AHGF.

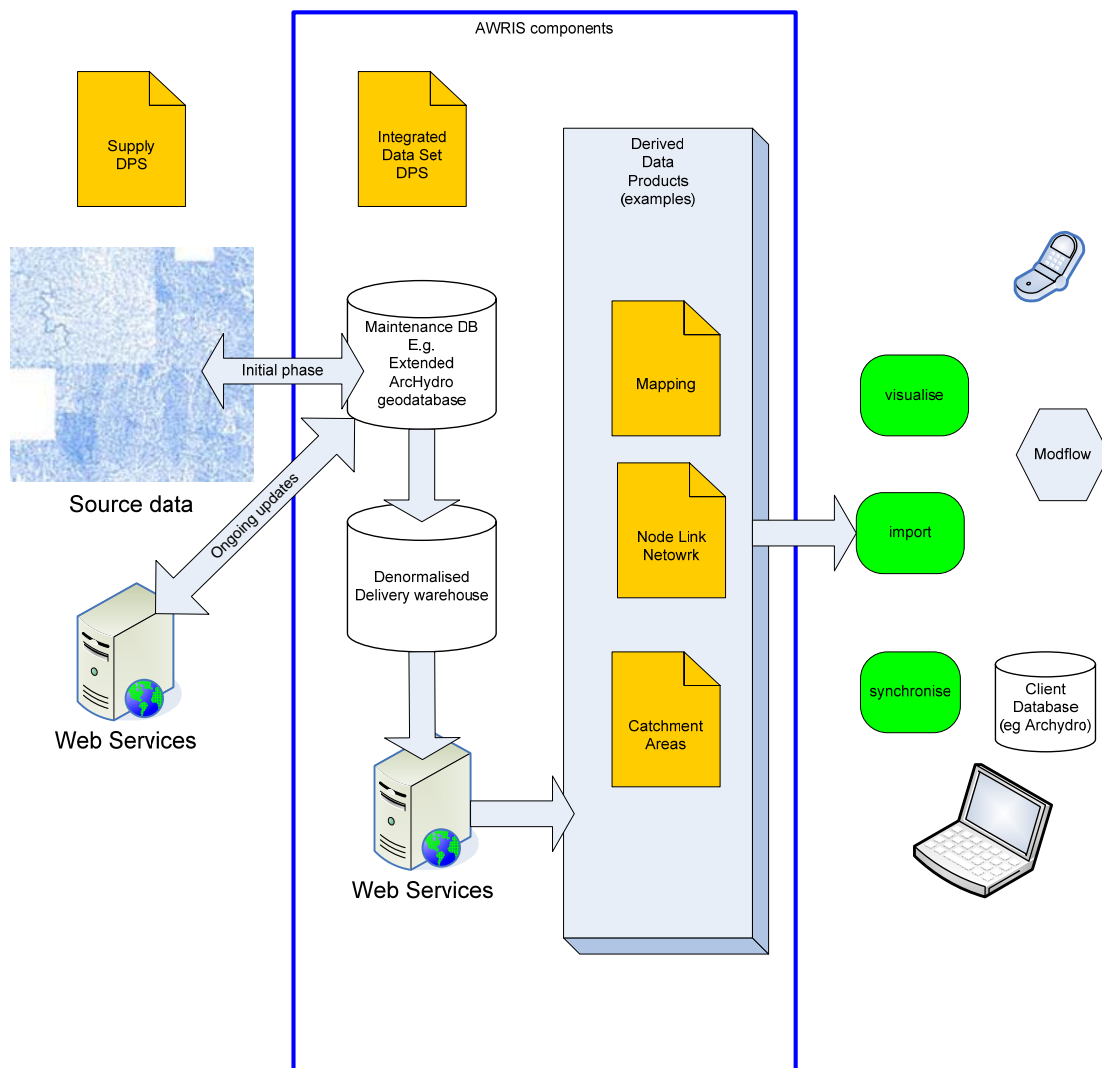


Figure 9 Distinction between supply and product delivery in the AHGF

From this architecture, it is important to note the following:

- initial data loading and ongoing updates may come from different sources. For example, initial loading may involve a process of collating and manipulating existing data sets. Updates may be received through a different mechanism. This shows the need to support evolution of data acquisition/harvesting techniques;
- the data maintenance system is separated from the distribution warehouse. This warehouse is structured for efficient delivery of data products and may be denormalised;
- output products will usually be vastly simplified views of the maintenance view of the data, optimised for specific applications; and
- there may be many output views. A critical subset must initially be identified for establishment in order to inform the design of data products;

## 5.2 Data Product Specifications

A key requirement of the AHGF is the need for a clear understanding between a data provider and a data subscriber on data product content. That is, the existence of a contract between both parties that describes such issues as:

- How the data product is to be identified;
- Who will own the data product;
- How the data product is structured;
- What quality criteria are to be applied to the data product; and
- How the data is to be delivered.

Each data product for the AHGF will require a robust specification of aspects of the data, from structure to data management processes. This Data Product Specification (DPS) should encompass the vast majority of metadata that would apply to the data set as a whole, any management units by which the data set is partitioned, services used to distribute the data set or support update transactions (operations), feature level and even attribute level metadata.

A DPS should allow creation of a per-instance metadata record that allows each feature to be fully assessed, and candidate updates generated from the field if required. In practice this will require per-feature metadata to be managed as part of the DPS, with specific feature attributes, or related metadata objects that refer to individual features, published as a capability of the data provisioning.

This in turn provides a straightforward design constraint on framework data sets.

ISO19131 provides the starting point for developing a DPS. A DPS will contain the following information:

- a) Overview – containing an informal description of the data product, abbreviations used, etc.;
- b) Specification scopes – describing where the product fits in a hierarchy of products (e.g. all AHGF data products will fit into an AHGF Scope) as well as spatial and temporal extent;
- c) Data product identification – a formal description of the data product including title, abstract, purpose and spatial resolution;
- d) Data content and structure – for feature based data sets, this section will include the feature type catalogue to be used. For coverage and imagery data sets, coverage identification information is given;
- e) Reference systems – the various reference systems (coordinate or other) to be used;
- f) Data quality – specifying data quality requirements;

- g) Data product delivery – describing the form in which the data is to be delivered including structure of files and versions of formats;
- h) Metadata – core metadata elements to be provided with the data set;
- i) Data capture (optional) – specifying data capture methods;
- j) Data maintenance (optional) – describing data maintenance processes to be implemented;
- k) Portrayal (optional) – describing how the data set is to be presented graphically; and
- l) Additional information.

DPSs will be required in a number of areas of the AHGF to specify the various input data sets as well as to describe derived data products.

In the case of input data sets, the DPS forms a key component of the contract between the BOM and a particular data provider. The DPS will describe what is required of the data set as provided by the data provider. It should be noted that this DPS may not necessarily describe the data set as maintained within the AHGF. It simply describes what the AHGF owner is expecting from the supplier.

In the case of DPSs for the various output data sets, a similar contract exists. However, in this case the agreement is between the AHGF and users of the output data. This agreement gives users confidence in the data.

There is a clear overlap between DPS and traditional data metadata statements. One important difference is that a DPS describes an intention. There is recognition that implementation realities may prevent exact delivery against the specification.

## 6. AHGF DATA SETS

As noted previously, the AHGF will consist of a collection of data sets that can be used as a national reference framework for hydrological modelling, forecasting, assessment and reporting. This section identifies an initial list of data sets that will be part of the AHGF.

### 6.1 Surface Water – Hydrology

The Surface Water – Hydrology component of the AHGF is a national representation of surface water in a drainage network. It consists of a collection of surface water features, for example streams, reaches, waterbodies, canals, lakes, reservoirs and wetlands, and the relationships between these features – the hydrological stream network (sometimes referred to as the ‘Blue Line Network’ as it corresponds to the blue lines depicted on maps indicating water features). Feature relationships are managed so that connectivity can be used for modelling purposes, such as flow tracing.

#### 6.1.1 Characteristics

There are a number of characteristics of the Surface Water – Hydrology component of the AHGF that need to be addressed. Some of these are described below.

- Water Features – a list of the surface water features that need to be modelled is required to ensure that all types of features are captured in the conceptual model. An initial list of features that should be included is: streams, floodplains, storages, wetlands, irrigation infrastructure and pipelines.
- Reaches – there is no single uniform description of a reach that suits all purposes, as the following list of definitions demonstrates:
  - Continuous extent, especially part of river that can be looked along at once between two bends, or part of canal between locks (Sykes 1984).
  - ...a significant segment of surface water that has similar hydrologic characteristics, such as a stretch of stream/river between two confluences, or a lake/pond. Reaches also are defined for unconnected (isolated) features, such as an isolated lake/pond (NHD 2007).
  - A length of river or stream, usually defined between confluences. Can also refer to a water body or a portion of a water body (Maidment 2002).

The definition of an AHGF reach needs to be adopted/defined. In the NHD definition, for example, the notion of ‘similar hydrologic characteristics’ would need to be established. If a reach is defined between confluences, then they could be hundreds of kilometres in length. This might be useful for some applications, but will not be for

others. In this case a reach would need to be composed of smaller segments, but this could be addressed as an implementation issue.

An example where ‘business rules’ are used to define reaches can be found in the National Land and Water Resources Audit I (NLWRA I). A reach was defined as having a minimum catchment drainage area of 50 square kilometres. Reaches used for the NLWRA I were formed by concatenating one or more network links and joined according to the following rules:

- all first order links (those with no tributary) were assigned as separate reaches;
- for links downstream, the product of link slope and drainage area (a stream power surrogate) was compared to that of the two tributary links;
- if one tributary link provided 90 per cent or more of the area of the link, then the tributary reach was continued downstream unless the area slope product differed by a factor of two or more. In that case a new reach was started;
- if no tributary link dominated the area of the downstream link, then the tributary reach of closest area and slope product was continued downstream. If the product of area and slope for the link differed by more than twofold from both tributaries a new reach was started; and
- new reaches were initiated at the entrance and exit from lakes and reservoirs (Norris et. al. 2007).

A reach is the building block of the stream network. There is a close connection between a reach and a drainage area catchment, discussed in Section 6.6 (Key Catchments), so the definition of a reach may need to factor in catchments also.

### 6.1.2 Node/Link Network

An alternative view of hydrological stream networks is that the building block is not a reach, but a ‘stream segment’ or ‘link’ whose purpose is solely to define the individual components of the stream network. A reach is defined on top of, or in terms of, the stream segments. A reach definition may not directly align with the junctions connecting the stream segments. For example a reach may be defined at an offset from a junction which is a significant location in terms of a specific modelling exercise.

This type of representation is referred to as a Node/Link Network. A node (junction) is defined where two or more streams meet. A link is the term given to the line connecting nodes. A node may be created on a link for different reasons: intersection with some other feature, for example a river goes under a bridge; a river crosses a political boundary (state/local government area); the node is the location of a monitoring point (gauge). It is not yet clear how these nodes impact the definition of a reach for a stream. They may demarcate a new collection of reaches or be defined as a feature on a reach.



### 6.1.3 Example Surface Water Hydrology

A first attempt at developing a conceptual model for Surface Water – Hydrology component of the AHGF is shown in Figure 10. This model shows both the surface water features (e.g. Reach and River) as well as the node/link representation of the network and how the two relate.

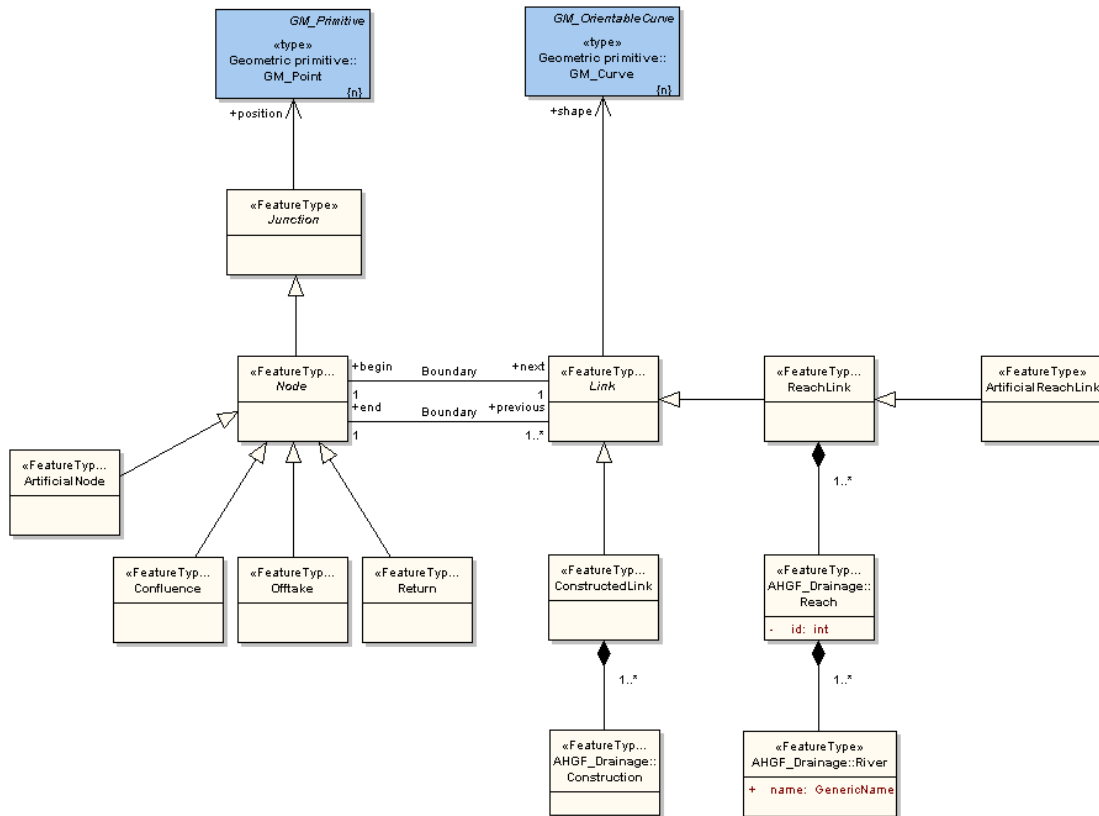


Figure 10 Example model of surface water - hydrology.

Figure 11 shows a related package of the model containing water body features. In this part of the model, waterbodies include both natural features such as lakes, and man-made features such as storage reservoirs that have been dammed.

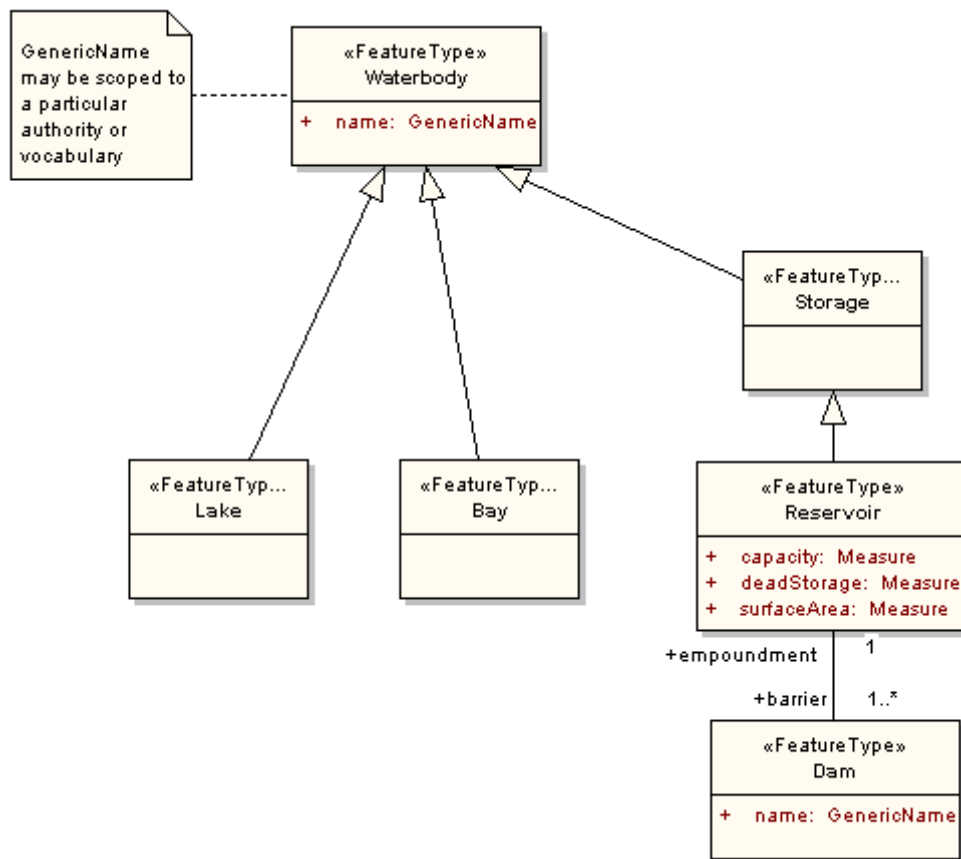


Figure 11 Waterbody features package.

## 6.2 Surface Water – Measurement Sites

In addition to the water features and stream network described in the Surface Water – Hydrology section above, the AHGF will need a nationally consistent data set for representing measurement sites (or monitoring points). Measurement sites may be the locations of stream gauges, metering points or off-take or return sites. Maintaining a consistent data set of these sites allows a consistent method of relating the associated measurements taken at these sites to be developed.

Figure 12 shows how measurement sites are represented within the draft conceptual model of the AHGF. This representation follows the OGC Observations and Measurements standard (Cox 2007a, Cox 2007b) where the Monitoring Point is associated with a feature of interest (for example the Dam in

Figure 12) or an element of a feature of interest (such as the Wellhead associated with a Bore in

Figure 12). The monitoring point remains independent of the actual physical feature, and all measurements and details of the monitoring process (e.g. the instrument or gauge used) are associated with the monitor point itself.

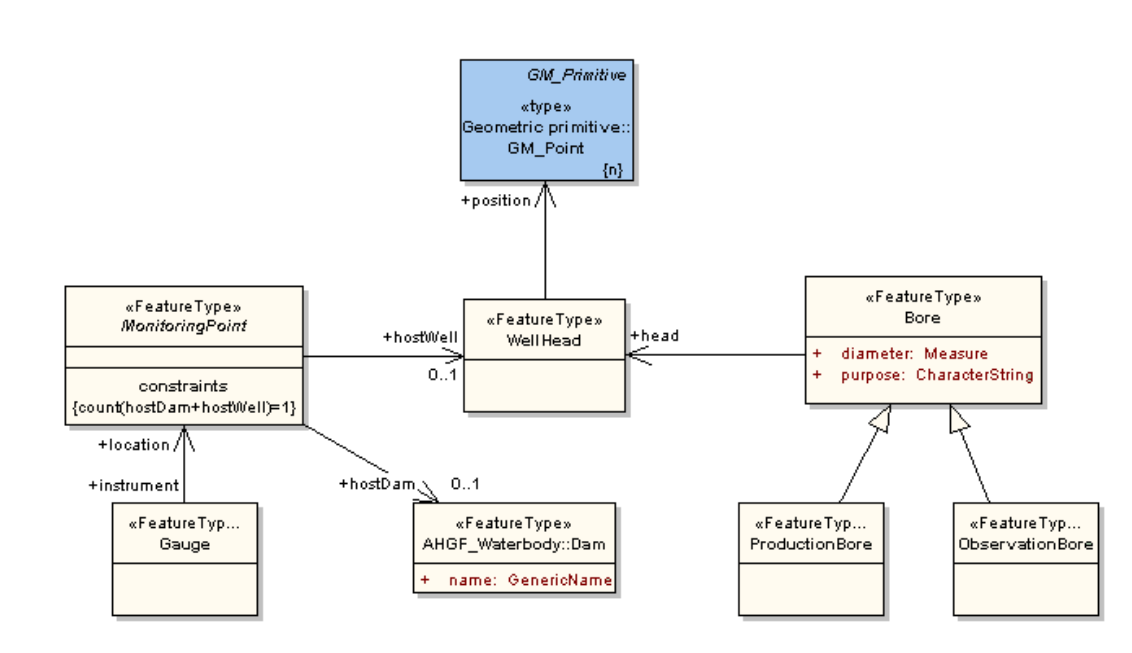


Figure 12 An example of how monitoring points are represented within the AHGF conceptual model.

### 6.3 Ground water – Measurement sites

Groundwater information provides an important component to Hydrologic studies. The main source of groundwater information comes from bore holes. Bore hole attributes may be collected when the hole is drilled. The bore location would be recorded along with other characteristics such as depth, purpose (stock & domestic, irrigation, public supply, observation), construction details and stratigraphy. This information categorises the different materials as layers. The hydraulic properties of the layers present are best measured in-situ but can be estimated from national data sets such as the Australian Soil Resource Information System (ASRIS) for soils (ASRIS 2008) and Groundwater Flow System (GFS) for aquifers.

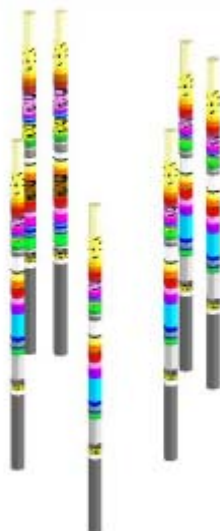


Figure 13 Example bore holes (Djokic 2008).

This description of bore hole data can be depicted as shown in Figure 13. The different colours represent the different layers identified when the bore hole was drilled. This information can be represented as a point location of the bore hole, the depth drilled, and the stratigraphy represented as material types with offsets to where a particular layer starts. A simple example is shown in Figure 14.

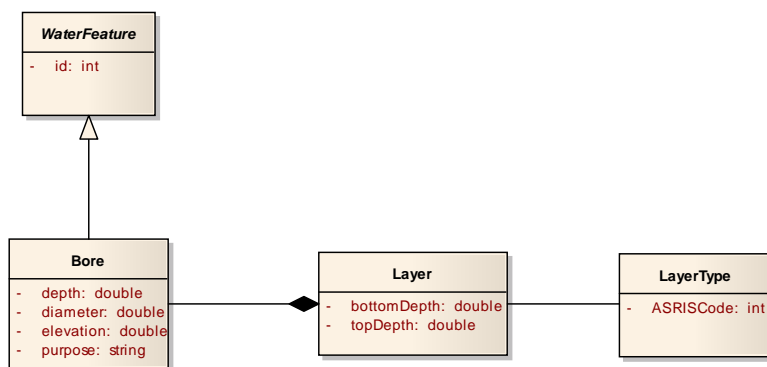


Figure 14 Bore hole model.

Note that Figure 14 has the Bore class inheriting from an abstract WaterFeature class. In the harmonised AHGF model, the Bore will be a primitive feature type and will provide an interface to the common AHGF identifier system (AHGF\_Element).

A distinguishing feature of bore hole data for groundwater studies is that the bore hole location will be three dimensional. As well as 'x'/'y' coordinates (latitude/longitude or similar) there will be a 'z' value, usually an elevation relative to some standard location. How this elevation attribute for a Bore is represented can be considered an implementation issue, or it can be a more fundamental characteristic of the model. Figure 14 uses an elevation attribute on the Bore class.

Once the bore hole is established, time series data can be collected measuring various parameters, such as the amount and rate of water extraction, water constituents ('hardness', salinity or other mineral content), water depth in bore hole or pressure (for capped artesian bores). This measurement data would be associated with a monitoring point as discussed above in Surface Water – Measurement Sites.

## 6.4 Ground Water – Hydrology

Hydrography concerns the surface water network and the environment, the land features through which water flows. Hydrology deals with the cycle of water circulation through the atmosphere, surface, subsurface, and their interactions. This cycle includes precipitation, surface runoff, groundwater infiltration, subsurface flow, evaporation, plant transpiration, and outflows (both from groundwater and surface water).

Hydrologic studies require more than a surface water stream network. Data needs to be collected from monitoring stations that record environmental observations of stream flow, groundwater levels, and rainfall. This information is combined with other data sets, such as geology, land use, soil characteristics, ground layer materials, and used as input to mathematical models. These models predict the flow of water through the landscape environment, and how the water properties change. Historic data can be used to calibrate the models to provide better predictions.

A data set of ground water features will be required within the AHGF. In particular, details of hydrogeologic features such as aquifers will need to be available to enable modelling. Geologic information can be used in conjunction with bore hole data to define the extent of an aquifer. Bore hole information is collected for a localised study for a small region. An aquifer covers a larger regional area and the aquifer is modelled more coarsely in terms of water inputs and outputs, when compared to information gathered about bore holes. An aquifer can be represented simply as a boundary on the terrain under which the aquifer exists, or it can be a more complex geological structure in which the various layers are defined in three dimensions.

While the initial conceptual model of the AHGF described in this document does not include ground water hydrology, insight into how this may be represented can be found in the Canadian Ground Water Markup Language and associated model (NRC 2008). This area will need further investigation.

## 6.5 Terrain

Digital Elevation Models (DEM) provide a representation of the three-dimensional nature of the landscape. DEMs are developed from point elevation data sets derived from measurements on-ground or through remote-sensing. A DEM can be useful for a range of activities, including hydrology where applications include derivation of catchments, stream networks and the slope of streams as well as for terrain analysis. An Australian, hydrologically consistent DEM will be an important data set for ensuring consistency, particularly in the use of derived products.

It is expected that the Australian Hydrological DEM will be developed from SRTM data at 1 second resolution to meet the needs of hydrological applications. Key issues that will need to be addressed in developing this data set include data quality standards and measures, what additional layers will be included with the elevation data (for example error surfaces) and delivery techniques.

## 6.6 Key Catchments

A nationally consistent collection of catchments is necessary to ensure valid comparisons can be made between data or reports, both from differing sources, and over time.

The definition of a catchment varies. In particular, a catchment in the AHGF context is not the same as a catchment as defined by ESRI, or as used in the HUC hierarchy. In ESRI and HUC a catchment is the drainage area in which water over the landscape flows to a well defined reach river segment. In Australia, a catchment refers to the entire upstream contributing drainage area with respect to an arbitrary drainage point – not just the local region.

This distinction is illustrated in Figure 15 where a simple river network is shown. Here the lower reach, defined between the two indicated nodes, has a catchment of the whole upstream network. By contrast, the ‘contributing area of a reach’ is just the lower drainage area.

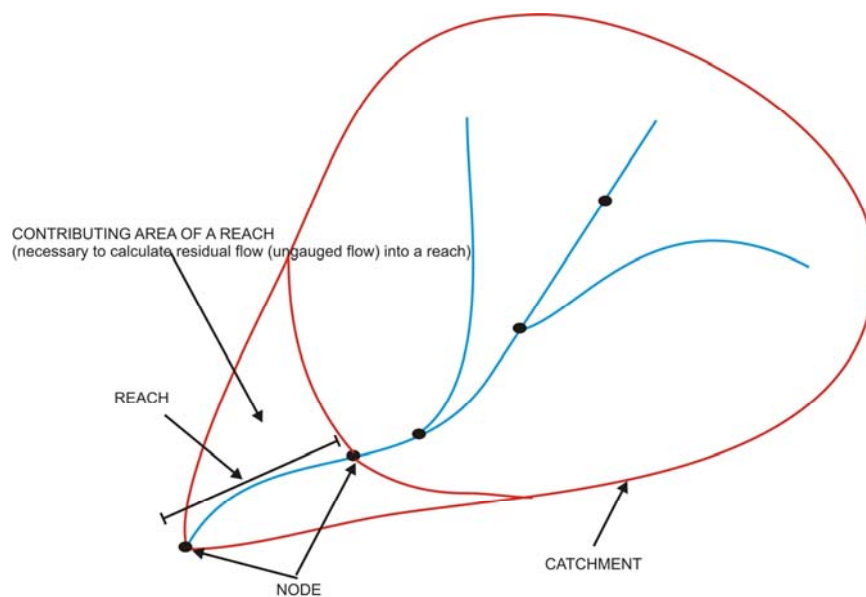


Figure 15 Demarcating drainage areas (Petheram 2008).

It should be noted that the alternative concept, of defining a drainage area corresponding to only the local region, not the whole region, still has a role to play in Australian water resources investigations.

While catchments have previously been defined in an arbitrary manner by different hydrologists, a standardised set of catchments will be defined for the AHGF and will be used as the basis for all water resource reporting and assessment. Where more or less specific regions



## 7. NEXT STEPS

The key to the success of the Australian Hydrological Geospatial Fabric (AHGF) will be:

- development of a modular conceptual information model, capable of supporting many use cases and associated implementations;
- consistent use of standardised Data Product Specifications describing clearly both data sets input to the AHGF and derived data products output from the AHGF; and
- adherence to an iterative development methodology that covers all aspects of the data product lifecycle and allows for evolution of technology, data products and user requirements.

This report has proposed a development methodology and a number of the components of a conceptual architecture for building, structuring and managing the AHGF. These proposals are based on a review of current best practice in delivery of similar products overlaid onto the known requirements of the AHGF.

A number of the methods proposed in this report are untested. Furthermore, some of the originally proposed content, in particular various pieces of conceptual models, have not been developed using the proposed methods. This was due to unavailability of suitable domain experts required to develop the key use-cases.

The key next steps for development of the AHGF are to identify and document these critical use cases. That is, the proposed methodology needs to be exercised.

It is also necessary to instigate the engagement aspects of this project initially planned but not undertaken. Early engagement with potential users of the AHGF will be vital if these users are to adopt it.

A number of research issues have also been identified whilst undertaking this project. In particular, the need for an identifier management scheme for the AHGF needs to be addressed as quickly as possible.



## REFERENCES

- Arc 2001, Hydro Data Model and Tools, 2001, ESRI, viewed March 2008, <<http://support.esri.com/index.cfm?fa=downloads.dataModels.filteredGateway&dmid=15>>.
- ASRIS 2008, Australian Soil Resource Information, CSIRO, viewed March 2008, <<http://www.asris.csiro.au>>.
- BD Carthage n.d., Base de Données sur la CARTographie THématique des AGences de l'eau et du ministère de l'Environnement, viewed April 2008, <[http://sandre.eaufrance.fr/rubrique.php?id\\_rubrique=17](http://sandre.eaufrance.fr/rubrique.php?id_rubrique=17)>
- BDRHF n.d., Base de Données Référentiel Hydrogéologique Français, viewed April 2008, <[http://sandre.eaufrance.fr/rubrique.php?id\\_rubrique=18](http://sandre.eaufrance.fr/rubrique.php?id_rubrique=18)>
- BRIDGE 2006, 'Chapter 3 Methodologies used for GWB delineation' BRIDGE Deliverable D10: Impact of hydrogeological conditions on pollutant behaviour in groundwater and related ecosystems, Volume 1 <[http://nfp-at.eionet.eu.int/Public/irc/eionet-circle/bridge/library?l=/deliverables/deliverable\\_d10&vm=detailed&sb=Title](http://nfp-at.eionet.eu.int/Public/irc/eionet-circle/bridge/library?l=/deliverables/deliverable_d10&vm=detailed&sb=Title)>
- BRIDGE 2005, Background cRiteria for the Identification of Groundwater thrEsholds, viewed April 2008, <<http://nfp-at.eionet.eu.int:8980/irc/eionet-circle/bridge/info/data/en/index.htm>>
- Cox, S. (editor) 2007a, *Observations and Measurements – Part 1 – Observation schema Version 1.0*, Open Geospatial Consortium Inc. Implementation Standard OGC 07-022r1, accessible from <<http://www.opengeospatial.org/standards/o%2526m>>.
- Cox, S. (editor) 2007b, *Observations and Measurements – Part 2 – Sampling Features Version 1.0*, Open Geospatial Consortium Inc. Implementation Standard OGC 07-002r3, accessible from <<http://www.opengeospatial.org/standards/o%2526m>>.
- Cripe, D & Grabs, W (editors) 2007, *The Integrated Global Water Cycle Observation (IGWCO) Theme – 3rd Annual Planning Meeting report*, 23 May 2007, available at <<http://www.wmo.int/pages/prog/igwco/reports/documents/igwco.plmtg.rpt.2007.igos14.pdf>>.
- Djokic, D. 2008, Hydrologic and Hydraulic Analyses Using ArcGIS, ESRI Course Notes, February 2008.
- EPA 2008, Total Maximum Daily Loads, United States Environmental Protection Agency, viewed April 2008, <<http://www.epa.gov/owow/tmdl/>>.
- IGN 2008, Institut Géographique National, viewed April, 2008, <<http://www.ign.fr>>.
- IHABBS 2005, Integrated Hydrologic Automated Basin Boundary System, National Weather Service, viewed April 2008, <[http://www.nohrsc.nws.gov/technology/gis/ihabbs\\_overview.html](http://www.nohrsc.nws.gov/technology/gis/ihabbs_overview.html)>.
- INSPIRE 2007, Drafting Team "Data Specifications" – deliverable D2.6: Methodology for the

## REFERENCES

- development of data specifications, INSPIRE Technical Report D2.6\_v2.0\_final.doc, available at <[http://www.ec-gis.org/inspire/reports/ImplementingRules/inspireDdataspecD2\\_6v2.0.pdf](http://www.ec-gis.org/inspire/reports/ImplementingRules/inspireDdataspecD2_6v2.0.pdf)>.
- Lefort, L. 2008, Towards defining a water data exchange format: State of the art review on water data standards, Technical report CSIRO ICT Centre 08/111.
- Maidment, D.R. (editor) 2002, Arc Hydro: GIS for Water Resources, Esri Press, 380 New York Street, Redland, California, October 2002, ISBN 1589480341.
- Maidment, D.R. & Tarboton, D.G. 2007, Data Sources for GIS in Water Resources Fall 2007 <<http://www.ce.utexas.edu/prof/MAIDMENT/giswr2007/visual/DataSources.ppt>>.
- National Groundwater Database 2008, National Resources Canada, viewed March 2008, <<http://ngwd-bdnes.cits.nrcan.gc.ca/service/ngwd/exploration/>>.
- NHD 2007, National Hydrography Data set (NHD), viewed March 2008, <<http://nhd.usgs.gov/>>.
- NHD 2006a, National Hydrography Data set Plus (NHDPlus), viewed March 2008, <<http://www.horizon-systems.com/nhdplus/>>.
- NHD 2006b, NHDPlus User Guide, viewed March 2008, <[http://www.horizon-systems.com/nhdplus/documentation/NHDPLUS\\_UserGuide.pdf](http://www.horizon-systems.com/nhdplus/documentation/NHDPLUS_UserGuide.pdf)>.
- NIDIS Team 2007, *The NIDIS Implementation Plan*, NIDIS Project Implementation Team, June 2007, available at <<http://www.drought.gov/imageserver/NIDIS/content/whatisnidis/NIDIS-IPFinal-June07.pdf>>.
- Normand M. & Mardhel, V. 2004, Mise en oeuvre de la DCE Comparaison entre les délimitations des entités hydrogéologiques de la BD RHF V2 et les masses d'eau souterraine (France métropolitaine), available at <<http://ftp.sandre.eaufrance.fr/public/incoming/bdrhfv2/RapportCompareBDRHfV2-MES.doc>>.
- Norris R., Dyer F., Hairsine P., Kennard M., Linke S., Read A., Robinson W., Ryan C., Wilkinson S. & Williams D. 2007, Assessment of River and Wetland Health: A Framework for Comparative Assessment of the Ecological Condition of Australian Rivers and Wetlands, available at <[http://www.water.gov.au/Publications/RWH\\_FARWH\\_Framework\\_070507.pdf](http://www.water.gov.au/Publications/RWH_FARWH_Framework_070507.pdf)>
- NRC 2008, Ground Water Markup Language Model, National Resources Canada, viewed April 2008, <<http://ngwd-bdnes.cits.nrcan.gc.ca/service/ngwd/exploration/ngwd/featureCollection.html?REQUEST=GETFEATURE&FEATUREID=ca.gc.nrcan.gwml.archivedocument.model&SESSION=&locale=en>>.
- OIEAU 2008, Office International de l'Eau, viewed April 2008, <<http://www.oieau.fr>>.

- OSTP 2007, A Strategy for Federal Science and Technology to Support Water Availability and Quality in the United States September, Office of Science and Technology Policy September 2007, available at <<http://www.sws.uiuc.edu/iswsdocs/wsp/SWAQ.PDF>>.
- Parashar, A. 2007, Australian Hydrologic Geospatial Fabric Project, AHGF Design Project Plan Outline, CSIRO Internal Report.
- Petheram, C. 2008, Personal Communication, March 2008.
- Power, R. 2007, Water Resources Data Models, CSIRO ICT Technical Report 07/xyz.
- SEE GRID 2007, *Hollow World: a GML application schema template*, Solid Earth and Environment GRID, viewed 1 April 2008, <<https://www.seegrid.csiro.au/twiki/bin/view/AppSchemas/HollowWorld>>.
- Smith, M., Koren, V., Zhang, Reed, S., Seo, D., Moreda, F., Kuzmin, V., Cui Z. & Anderson R. 2004, NOAA NWS Distributed Hydrologic Modeling Research and Development, NOAA Technical Report NWS 45, available at <[http://www.weather.gov/ohd\\_files/data/hydrology/Distributed%20Modeling%20Review%20ver%20%20NRC%20Read%20FirstFekadu.doc](http://www.weather.gov/ohd_files/data/hydrology/Distributed%20Modeling%20Review%20ver%20%20NRC%20Read%20FirstFekadu.doc)>.
- Strassberg G. 2005, A Geographic Data Model for Groundwater Systems, PhD Thesis.
- Sykes J.B. (editor) 1984, *The Concise Oxford Dictionary*, Oxford University Press, Walton Street, Oxford, ISBN 0-19-861131-5.
- US GAO 2003, *Freshwater Supply: States' View of How Federal Agencies Could Help Them Meet the Challenges of Expected Shortages*, United States General Accounting Office GAO-03-514 July 9, 2003, available at <<http://www.gao.gov/new.items/d03514.pdf>>.
- USGS 2008a, National Water-Quality Assessment Program (NAWQA), United States Geological Survey, viewed April 2008, <<http://water.usgs.gov/nawqa/>>.
- USGS 2008b, SPATIally Referenced Regressions On Watershed Attributes Model (SPARROW), United States Geological Survey, viewed April 2008, <<http://water.usgs.gov/nawqa/sparrow/>>.
- Vogt, J.V., Soille, P., de Jager, A., Rimavičiūtė, E., Mehl, W., Foisneau, S., Bodis, K., Dusart, J., Paracchini, M.L., Haastrup, P. & Bamps, C. 2007a, A pan-European River and Catchment Database. EC-JRC (Report EUR 22920 EN) Luxembourg, 120 p. <[http://agrienv.jrc.it/publications/pdfs/CCM2-Report\\_EUR-22920-EN\\_2007\\_STD.pdf](http://agrienv.jrc.it/publications/pdfs/CCM2-Report_EUR-22920-EN_2007_STD.pdf)>.
- Vogt, J.V., Soille, P., de Jager, A., Rimavičiūtė, E., Mehl, W., Haastrup, P., Paracchini, M.L., Dusart, J., Bodis, K., Foisneau, S., Bamps, C., 2007b, Developing a pan-European Data Base of Drainage Networks and Catchment Boundaries from a 100 Metre DEM, Proceedings AGILE International Conference, May 2007.



### Contact Us

Phone: 1300 363 400

+61 3 9545 2176

Email: [enquiries@csiro.au](mailto:enquiries@csiro.au)

Web: [www.csiro.au](http://www.csiro.au)

### Your CSIRO

Australia is founding its future on science and innovation. Its national science agency, CSIRO, is a powerhouse of ideas, technologies and skills for building prosperity, growth, health and sustainability. It serves governments, industries, business and communities across the nation.