AuScope: Australian Earth Science Research Spatial Information Infrastructure

Dr Robert WOODCOCK, Australia

Key words: Education; Geoinformation/GI; GPS; Professional practice; Standards

SUMMARY

In 2006 the Australian Government announced a new funding initiative, the National Collaborative Research Infrastructure Strategy (NCRIS). NCRIS aims to provide Australian researchers with access to major research facilities, supporting infrastructure and networks necessary for world-class research. As a component of this strategy, $42.8 million was allocated to the Australian earth science research community to build an integrated national geoscience infrastructure system called AuScope.

AuScope will provide a number of infrastructure components to assist in understanding the structure and evolution of the Australian continent. These include the acquisition of subsurface earth imaging and earth composition and age analysis, a virtual drill core library, geological process simulation, and a high resolution geospatial reference framework.

To draw together information from these new initiatives and from other existing sources in academia, industry and government, AuScope is developing a world-leading geoinformatics network. This Community Earth Model will use open geospatial standards to allow real time access to data, information and knowledge stored in distributed repositories. It will be built on 'end-to-end' science principles whereby there will be access to the highly processed information and knowledge as well as the original raw data and the processing programs used to generate the results.

The key to linking resources in the Community Earth Model will be web service based access to the geoscience information holdings and computational services using common service interface and information models, including the emerging international standard for geoscience information, GeoSciML, and the Australian developed standard for mineral occurrences. The AuScope Portal provides the earth science community with a 'Google-style' interface to the Community Earth Model data and services.

AuScope: Australian Earth Science Research Spatial Information Infrastructure

Dr Robert WOODCOCK, Australia

1. INTRODUCTION

Many major earth science challenges of today involve complex system interactions. Industry, government and academic researchers now draw on information and analysis techniques across multiple disciplines and combine these to produce models of systems behaviour in an effort to answer questions of government policy, management or knowledge discovery. As a result of this interaction, there is an increase in data use outside of the discipline or organisation that would traditionally manage it. In addition, these science challenges must often be tackled at national or global scale in order to produce valid results (e.g. weather and climate prediction). The result is an increasing need to combine information gathered by different organisations responsible for different geographic regions.

Consider for example data collected by Federal, State and Territory Geological Surveys on the Australian Continent. A traditional user of this information has been the minerals industry. The minerals industry would combine data from multiple jurisdictions to target exploration nationally. By itself this created issues associated with integrating data across the nation, a problem clearly articulated in the Australian Government's Minerals Exploration Action Agenda – the Road to Discovery (DITR, 2004).

The Australian Continent is the platform on which we and most flora and fauna live. Its soils are derived from the rock base, most of our water resources reside within it as groundwater, and it is the storehouse of future clean energy, as well as a potential sink for the green house gas emissions. Hence, many of today's research challenges are geoscience related and data collected by Geological Surveys are used in problems such as climate change, sustainable exploitation of energy, mineral and water resources, predicting living with extreme geological activity, and managing disaster reduction. Data providers now have to service a wider range of disciplines and end users and the traditional hard copy maps, or more recently digital maps available via web sites, must evolve to meet the demands of integrated, global, multidisciplinary research efforts (Wyborn, 2008).

These challenges are not unique to Geological Surveys or geology. Many disciplines and many organisations are now struggling with data integration and distribution issues. Global or continental scale compilations can take years of effort and results may conflict with data available from other sources. Organisations publishing data are faced with a myriad of data formats to publish and, worse yet, multiple vocabularies describing that information. These interactions present significant operational challenges for researchers and service providers. Researchers are often faced with the choice of simplifying the research agenda or simply not undertaking the research at all.

A paradigm shift is required to resolve these operational challenges, and interoperability is that new paradigm. An interoperable national spatial data infrastructure (SDI) providing discovery and exchange services based on open standards, connecting distributed data sources and processing services hosted by industry, government and academic organisations is required. In 2006 the Australian Government through the Department of Industry, Innovation, Science and Research announced a new funding initiative, the National Collaborative Research Infrastructure Strategy (NCRIS). NCRIS aims to provide Australian researchers with access to major research facilities, supporting infrastructure and networks necessary for world-class research. As a component of this strategy, $42.8 million was allocated to the Australian earth science research community to build an integrated national geoscience infrastructure system called AuScope.

AuScope (www.auscope.org.au) is establishing a national SDI in a collaboration involving industry, government and academic organisations. AuScope consists of several components (Figure 1): Earth Imaging, Composition and Age, Virtual Core Library and Geospatial, are principally about new data acquisition to enable an increasingly clear and rich picture of the subsurface to be created. They will result in the acquisition and publication of seismic, magneto-telluric, geochemical, and hyperspectral core logging data and products. All of this data will be underpinned by an enhanced geospatial reference system from the Geospatial component.

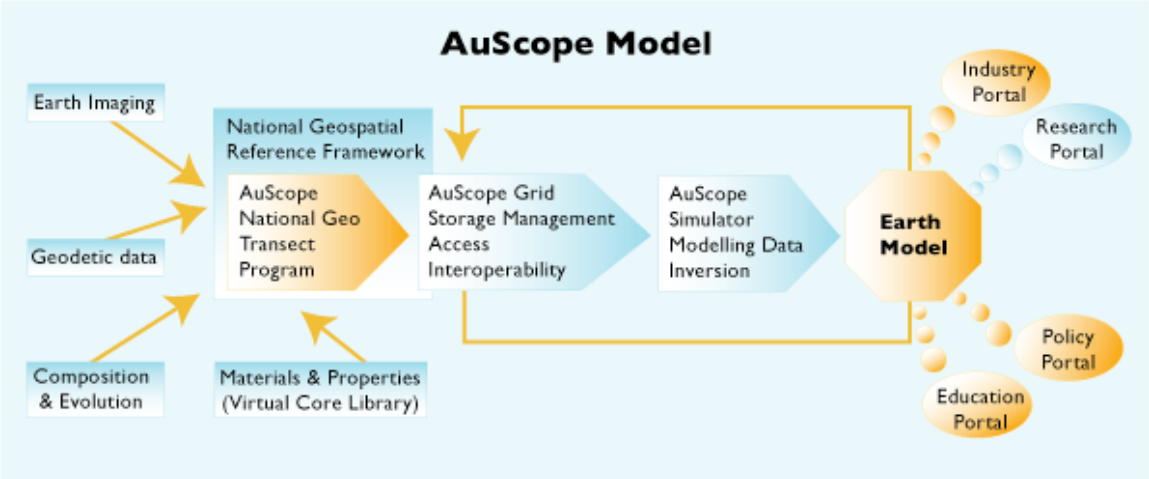

Figure 1

To draw together information from this new national infrastructure and from other existing sources in academia, industry and government, AuScope has developed a world-leading earth science research spatial data infrastructure – the AuScope Grid (www.auscope.org.au/category.php?id=10). This network provides access to independently managed data and computing facilities distributed around Australia. The network uses open geospatial consortium standards (OGC) and GeosciML (www.geosciml.org) to allow real time access to data, information and knowledge stored in distributed repositories hosted by different organisations. All of this information and the services are made accessible via the AuScope Discovery Portal (portal.auscope.org). An open source Spatial Information Services

Stack is freely available to allow organisations to add data holdings and computational services to the network. Any standards compliant middleware will operate and other choices do exist both commercially and in open source forms.

The use of open standards means researchers are able to build their own client access tools to pull together information, perform analysis and publish their results – scientific workflow "mash-ups". As more researchers and organisations do this the amount of accessible information and tools increases allowing for service chains to be built with simplified interactions between collaborating parties. Computationally demanding geoscience programs, ranging from geodetic modelling, earthquake and tsunami simulation through to ore formation and block caving, can be used to process observational information created by one party and as a basis for policy or business management decisions at the far end of a decision support service chain.

A key to success is achieving sufficient participation and open access to enable a thriving community of practice to develop and use the infrastructure. The Web based informatics infrastructure for AuScope is largely in place. Now, in research centres, laboratories, libraries and agencies around Australia, a wide variety of geological survey records, imagery, maps, supercomputers and sophisticated software services are online or coming online to fulfil the infrastructure's potential. The systems are linked as a "system of systems", through high bandwidth networks and they communicate and interoperate across the Web through standards-based interfaces and encodings. As resources come online, each becomes a working part of a vast meta-resource unified through the standards-based spatial data infrastructure.

2. THE AUSCOPE MODEL

The AuScope Model includes data acquisition, integration, and discovery and analysis components. It is essential that the various components in the infrastructure can operate as a seamless workflow. This workflow crosses organizational and jurisdictional boundaries and embraces the real issues of research collaboration. It provides direct support to multi-disciplinary and multi-organization research.

2.1 The AuScope Discovery Portal

The Discovery portal (Figure 2) provides a web based interface for search and accessing data, information, services and applications connected to the AuScope infrastructure. It allows users to discovery, browse, save and process geospatial information from earth science data sources around Australia.

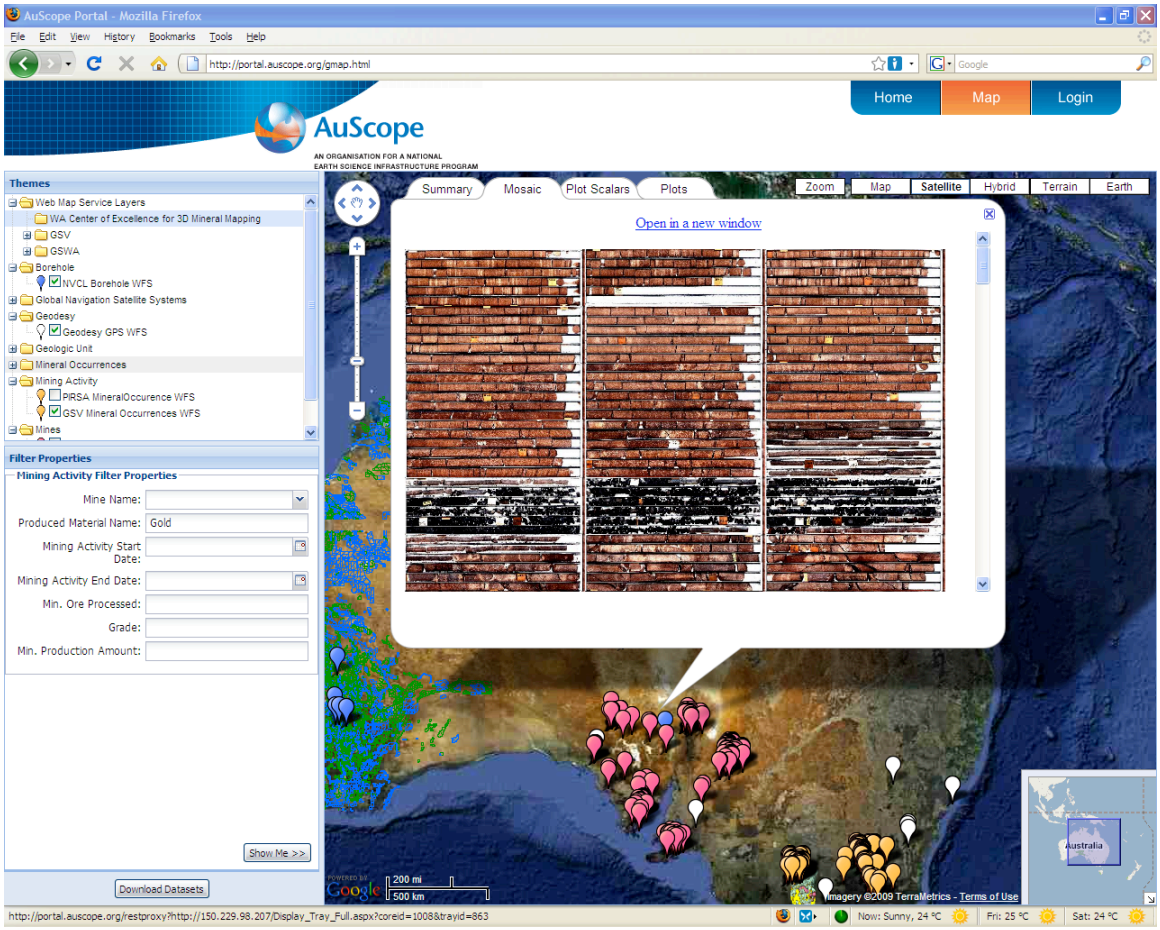

Figure 3

The research portal, and the underlying services, provides access to data using GeoSciML which was developed through the interoperability working group of the Commission for the Management and Application of Geoscience Information (CGI), a commission of the International Union of Geological Sciences (IUGS). The working group consists of geology and information technology specialists from agencies in North America, Europe, Asia and Australia.

Using GeoSciML, agencies can exchange data between organizations that are using different database implementations and software environments. GeoSciML will also facilitate access to geosciences applications available as OGC web services.

The portal allows users to select and display a variety of data types from a variety of service providers. Hyperspectral, borehole, global navigation satellite, geodesy, mineral occurrence and geology data are now available through the portal. The user can view data, filter the data based on user specified queries and download or deliver data to desktop applications.

Third-parties can make their data available via the Portal and need only deploy standards based middleware to support GeoSciML and other Open Geospatial Consortium standards

like Web Feature Services and Web Map Services. Once those services are registered with AuScope they will appear in the portal. Additionally, third parties can discover and access the underlying services the portal uses. This allows custom portals and workflows to be built which cross discipline boundaries (e.g. combining ground water information with geology, minerals chemistry and environmental information).

2.2 The Spatial Information Services Stack

The Spatial Information Services Stack (SISS) is a collection of open source projects which have been modified to operate together and provide a complete spatial data interoperability package including WFS, WMS, GML application schema development, registries and vocabulary services. The AuScope Grid project is deploying the SISS across multiple Australian Government agencies and research organizations, assisting those organizations in making the transition to supporting a national spatial data infrastructure as a routine operational service.

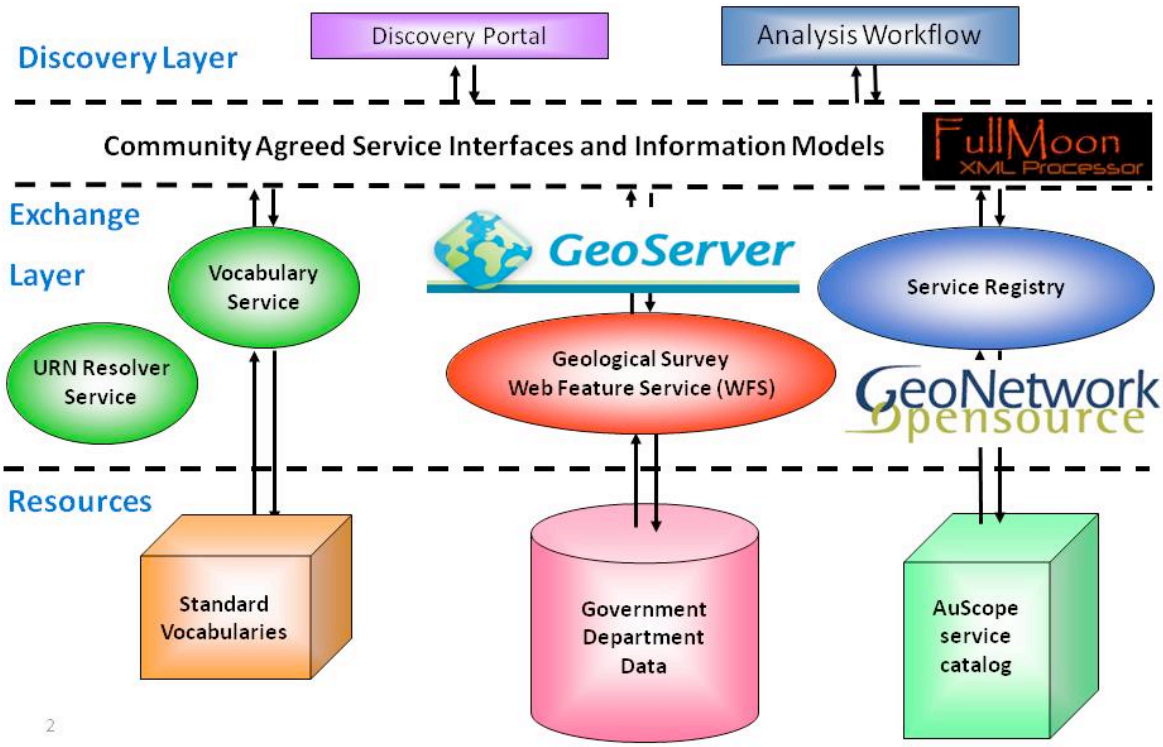

Figure 4

Whilst initial interest has been in the geosciences, and in particular the State and Territory Geologic Surveys, the SISS is quite generic. Additional use has been seen in ground water and water related agencies using WaterML for data exchange. Recently, the CSIRO, the lead agency for the AuScope Grid initiative, has been awarded additional funding through the NCRIS Australian National Data Service to take the SISS and its community building approach to other agencies, beyond the geosciences. This project, called the Australian Spatial Research Data Commons, will build on the SISS activity and broaden the reach of the spatial

data infrastructure and the supported research domains. In doing so, the spatial data infrastructure being developed by individual NCRIS capability areas and other organizations will be able to interoperate and thus share resources and establish a spatial information data commons that will also be accessible by other communities of interest.

3. CONCLUSION

The research infrastructure described above makes it possible to federate nationally distributed data sets, to develop tools to manipulate large data volumes and to establish an appropriate governance framework to ensure sustainability. AuScope Grid will be augmented with distributed data storage hardware, high bandwidth network links, data management protocols, middleware and software, all of which will enable AuScope to be substantially more than the sum of its parts. A key challenge will be to link the major geoscience and geospatial data stores of the government agencies with the high performance computing resources and high bandwidth networks of the academic community and on into industry.

A coordinated approach to data acquisition, analysis and simulation and modeling within the geology community by itself is not sufficient. There has to be cross-community communication to enable integration of the new AuScope compute and data grids with those of other research communities (Water, Environment, Spatial, etc.). To achieve this, the Australian Spatial Research Data Commons project will work with other discipline to broaden the skill base and deployments of the SISS by providing training and support to groups interested in publishing their data as contribution to the research data commons.

AuScope is leading a major transition in data delivery for geological organizations and researchers. Geoscience data is becoming more easily accessible, and the latest versions of that data can be dynamically accessed. Growth in the size of the network increases the value of the data and the value of tools and web services used to process the data. Because other communities of interest are building similar systems that are interoperable with AuScope the data is becoming increasingly useful in cross-disciplinary efforts to tackle transnational issues such as climate change and resource utilization.

The Auscope spatial services infrastructure is based on open source technologies and is available freely. For further information contact Robert Woodcock (Robert.Woodcock@csiro.au).

REFERENCES

DITR (Australian Department of Industry Tourism and Resources), 2004, Minerals exploration : the road to discovery : the minerals exploration action agenda

Wyborn, L., Cox, S, Woodcock, R, 2008, OneGeology: assisting Geological Surveys worldwide to interoperate seamlessly on the next generation internet, The 33rd International Geological Congress

BIOGRAPHICAL NOTES

Dr Robert Woodcock is Stream Leader for the CSIRO Minerals Down Under Flagship's National Geoscience Data Infrastructure Stream. He is responsible for the Auscope Grid and National Virtual Core Library projects which are deploying an OGC based Australian geoscience spatial data infrastructure. He also manages the flagships Solid Earth and Environment Grid community website which hosts the GeoSciML OGC application schema community. Robert has some 17 years IT R&D experience in industry and government research organisations. His skills span collaborative project leadership and development, software engineering, and spatial information management, visualisation and simulation and analysis services.

CONTACTS

Dr Robert Woodcock
CSIRO
26 Dick Perry Ave
Kensington, WA, 6151
AUSTRALIA
Tel. +61 8 6436 8780
Fax +61 8 6436 8555
Email: Robert.Woodcock@csiro.au
Web site: www.csiro.au