Bringing information management practices to natural disaster risk reduction

Monica Osuchowski argues that the concepts behind the existing multi-organisational virtual information database on landslides can be applied to the all hazards environment to provide sound hazard knowledge and disaster risk reduction.

Abstract

The important role of information management in improving baseline data for natural hazards has been demonstrated through a collaborative pilot project between Geoscience Australia, Mineral Resources Tasmania and the University of Wollongong. The result is a 'virtual' landslide database that makes full use of diverse data across three levels of government and has enabled landslide data to be collated and accessed from a single source.

Such a system establishes the foundation for a very powerful and coordinated information resource in Australia and provides a suitable basis for greater investment in data collection. This paper highlights the capacity to extend the methodology across all hazards and describes one solution in facilitating a sound knowledge base on natural disasters and disaster risk reduction.

Introduction

It is generally acknowledged that effective disaster risk reduction requires a systematic understanding of the history of natural hazard events. At the core of this lies a fundamental need for data as acknowledged in the Council of Australian Government (COAG) report on natural disasters. The report through Reform Commitment 2 (RC2) called for the establishment of a 'nationally consistent system of data collection, research and analysis to ensure a sound knowledge base on natural disasters and disaster mitigation' (COAG 2004).

Developing consistent data across a single hazard is challenging enough, but developing consistency across a broad range of hazards is significantly more complex. Recent advances in information management methodologies have provided the opportunity to pursue a new approach in data management, which has the capability to meet RC2. The approach utilises interoperability techniques and was successfully tested and implemented in a pilot project to facilitate consistent landslide data.

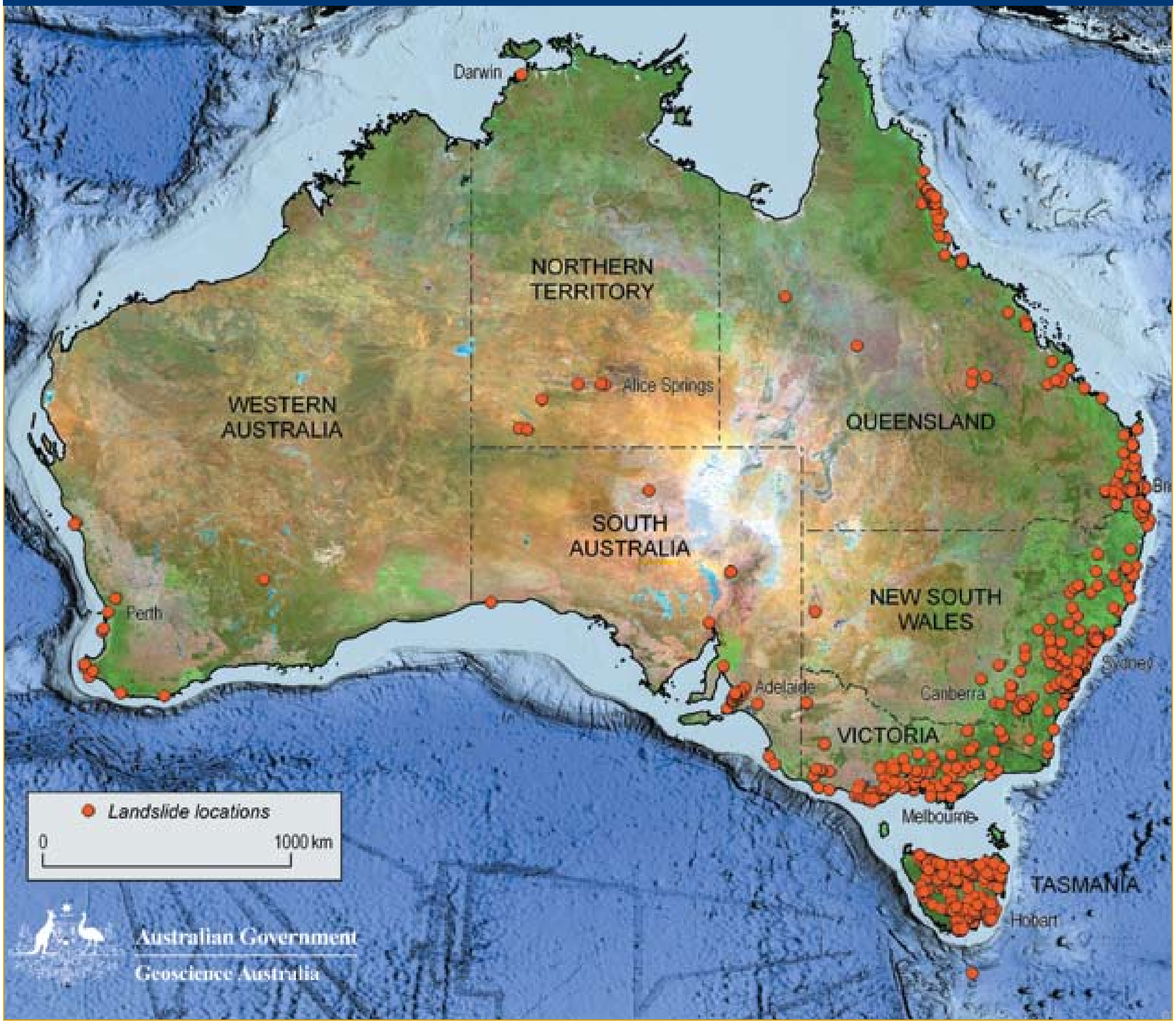
Drivers for coordinated landslide data

Despite the frequent and ongoing occurrence of landslides across the most populated regions of the Australian coastline (Figure 1) the cost of landslides in Australia is unknown. It is believed the annual average cumulative cost may be comparable to other higher profile natural hazards. However, challenges in data collection and the absence of cost measures commonly used for reporting on cyclone, hail or bushfire for example, such as either the amount of insured loss or relief funding, means it is difficult to estimate the cost of landslides. A single landslide event rarely meets the threshold levels required for relief funding, and insurance for landslide damage is not provided. This means costs are absorbed directly by the local government, private home-owners or infrastructure authorities.

Capturing landslide data and making this information available to those who need it was identified as an underpinning requirement in susceptibility, hazard and risk mapping and also for risk analyses, research and land-use decisions (AGS, 2007).

Therefore, improving our collective knowledge of landslides in Australia is essential.

Figure 1. The distribution of landslide events in Australia recorded since 1842.



Challenges in landslide data collection

Landslides are perhaps one of the most difficult hazards in terms of obtaining and collating data due to the localised responsibility of individual impacts. This means there are a wide variety of approaches that individuals use in managing information, and subsequently data generally is:

- widely dispersed;
- in different formats;
- of varying levels of detail;
- difficult to access; or
- not reported.

Implications are that data cannot readily be collated across different sources, compared or aggregated. This presents difficulties to others needing access to information for decision making, such as geotechnical practitioners or other levels of government.

There are two conventional solutions for achieving consistency in data collection:

1. responsibility falls directly to a single organisation; or,
2. responsibility is shared by everyone collecting data to an agreed standard.

However, the challenges in developing consistent landslide data collection using the aforementioned solutions lie within the following:

- while Geoscience Australia (GA) maintains a national landslide database in an internally consistent format, it only captures those events reported in the media. Consequently, the database severely underestimates the true occurrence of landslides and this is shown in Figure 2;
- trying to physically incorporate landslide data from a large number of sources and maintain it in a central database is impractical and resource intensive given the diverse approaches utilised;

- in considering the nature of landslide occurrence and the size of the country, it would be expensive and inefficient for a single agency to collect data in a consistent way that was useful to those needing it across all levels; and
- imposing a standard for a consistent approach among many individuals is also not feasible due to the number of individuals collecting data and because their existing data capture systems meet their needs.

Figure 2. Comparison of the number of landslides available in the national database managed by GA in comparison to the actual number of landslides available at local scale.

Locality	GA database	Local database
Tasmania	69	2074
Wollongong	72	402
Pittwater Council	6	193
TOTAL	147	2669

Our aim

GA's aim, as a technical advisor in the implementation of RC2, was to find a way to achieve national consistency in data collection while acknowledging existing data collection efforts. Due to some of the aforementioned challenges in landslide data capture, it was important to think beyond traditional solutions and consider innovative alternatives. A vision was needed to encapsulate what the most efficient way would be to collect and manage data and what the future of data management might look like.

The vision

An effective way of managing and utilising landslide and other natural hazard datasets across all levels (eg: local, regional and national levels) is embedded in a few simple concepts:

- it should be possible to collect data once and maintain it at the most effective and appropriate level;
- it should be possible to combine spatial information and share it between many users and applications; and
- it should be possible for information at one level to be shared at all levels.

These concepts are analogous to several of the stated visions of the INSPIRE initiative underway in the European Union (INSPIRE, 2008).

A solution

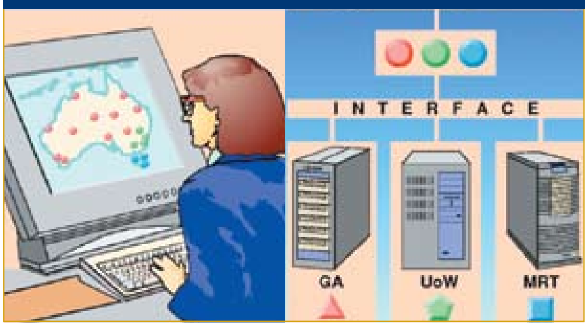
An information management methodology known as "network service-oriented interoperability" was identified by GA as one solution to overcome the challenges described across data capture within the landslide domain.

Interoperability, in the way that GA decided to implement it, acts like an information portal. The idea is that information located in physically separate databases can be viewed through a portal as one consistent virtual dataset. The virtualisation is achieved through the ability to collate and characterise large volumes of information over the internet regardless of how individual database custodians decided to manage and describe their data. It does this through mapping or translating unique data into a common format via a web interface. This interface essentially acts like a buffer between a user searching for data and each database provider, translating information back and forth as required (Figure 3).

Implementing an interoperable approach by using available databases as they means that existing data collection efforts are acknowledged and that full value is made of captured data. It is important to emphasise that database custodians retain complete responsibility for their own data. Each continues to collect, manage and maintain data as they always have, and in which ever way best meets their needs.

This means it is possible to collate a variety of data from different organisations without imposing change on individuals or agencies (i.e. developing consistency using a 'bottom-up' approach). The outcome is that such data not only continues to serve the needs of individual database custodians, but also serves a broader need.

Figure 3. Concept underlying the common interface into one 'virtual' database.



A pilot project

GA worked in partnership with Mineral Resources Tasmania (MRT) and the University of Wollongong (UoW) to demonstrate a way of establishing consistency across national, regional and local scale landslide data and to showcase some of the benefits and functionality of adopting such an approach. This pilot project is referred to as the Landslide Database Interoperability Project (LDIP). Each landslide database forming part of the LDIP contained different amounts of data, expressed details differently and was created in a different format including Oracle (GA), Microsoft Access (MRT) and Microsoft Excel (UoW).

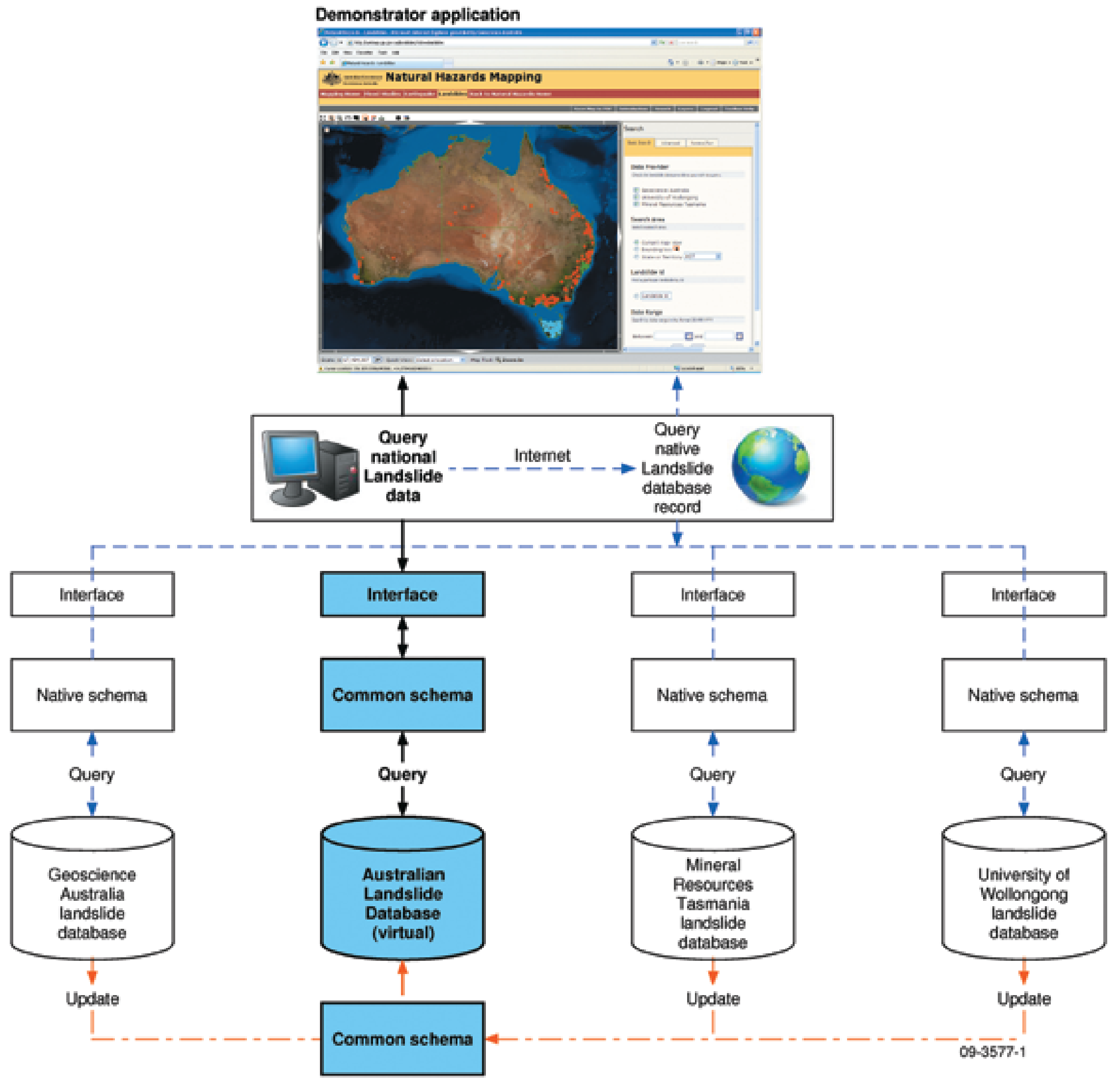
The LDIP sought to gain experience in applying new techniques and ascertaining their effectiveness as a way of potentially meeting RC2 for all hazards. Technical components were developed between Social Change Online, CSIRO and GA's Information Services Branch. The project was explicitly designed to exercise and consolidate an emerging methodology for designing such data services. Therefore, many important aspects required for ongoing sustainable use were beyond the scope of the pilot.

It is important to emphasise that the LDIP does not encompass all of the "data collection, research and analysis" issues which need to be addressed under RC2. However, it provides a simple means to highlight the complexity of data and information management for natural disaster mitigation and provides a new perspective in the way such challenges can be overcome.

Key to an all-hazard approach

The key to the all hazard approach adopted by GA were the strategic decisions to adopt common vocabularies and establish the system upon a common conceptual data model. The significance of how and why these components were established and the importance in relation to extending the approach across a range of natural hazards are the focus here. The project methodology is described in Osuchowski & Atkinson (2008).

Figure 4. A conceptual overview of the interoperable database.



Common vocabularies

An application schema is a set of definitions which describes how data is structured and expressed. It determines how data is related to other domains such as rainfall or geology. It also describes how a user will search and query data, and the way results are presented to them. As such, the schema forms a crucial part of the 'interface' alluded to earlier.

In order to create an application schema for landslides and thereby present diverse landslide data in a consistent way to users, many specific landslide models (or native schemas) needed to be synthesised into one common 'rich' schema (Figure 4).

This synthesis was achieved through reaching agreement within the science community on a set of common vocabularies to describe landslide events. It was necessary to find common ground for describing analogous information. An example can be shown in that all landslide databases typically capture information about the cause of landslides, but each has its own way of describing this (e.g.: rainfall events, precipitation, flood conditions, blocked drainage, fill failure, weak materials, excessive loading). Landslide causes are limited and it is possible to agree on what these causes are. For example we can separate 'cause' into contributing and/or trigger factors, which are either natural or man-made. Natural factors can be broken down into themes like 'ground conditions', 'geomorphological' or 'physical' with a series of terms used to provide more detail within each theme. In many cases where international conventions were available they were adopted more explicitly. Popescu (1994) was adopted to describe the cause of landslides in the landslide application schema. It is important to reiterate here that each database custodian retains own original data descriptions (native schemas) and the common schema referred to here simply is a veneer overlaid upon each database which maps data into the common format via the web.

As part of developing an application schema, it was important to be conscious of the different users of landslide data and the type of information they need, because the way in which information is recorded, has implications for how useful it is to users.

A common conceptual data model

The common vocabularies are a key part of the data model used for the interface. The data model contains the instructions for the transfer and exchange of data. For an all hazards approach, a common conceptual data model and the use of standards were essential. These are what can ultimately enable data to be collated and shared across multiple natural hazard databases in the future.

Consider for example the nature of landslides and landslide investigations. Landslides typically have a strong geospatial component and, as a result, landslide data is often displayed and managed with databases and GIS technology. It is important to realise that these components are not specific to landslide databases, but are also true for other natural hazards, and are in fact also generic with regards to the way any spatial data is captured. Therefore, it is efficient to leverage off international developments in geospatial standards which define how this data is exchanged (Cox & Richard, 2005). By doing this, it provides us with the ability to directly link and incorporate data from related domains as they are developed in future. For example, we could query relationships between landslide data with detailed datasets on earthquake, rainfall, soil, geomorphology and geology, which could further aid more consistent susceptibility, hazard and risk assessments.

In many cases the type of information described or required in landslide inventories is also analogous to information described or required in other natural hazard databases. Consider for example the damage following an event such as number of buildings damaged or destroyed, type of direct or indirect damage, remediation costs, etc. It is important to be able to collate this type of data across all hazards. Therefore, it makes sense for a generic damage/impact model to be developed and applied across all hazards in future. Customisations for specific hazards if they are required could be undertaken from this common point. Adoption of such an approach would allow for information to be easily aggregated across all hazard databases (or all other domains that deal with a component of damage, such as biological or technological hazards).

To reflect such possibilities, the landslide model was developed in 'packages' (a way of compartmentalising information) so that an individual package such as 'damage' for example, can be easily extracted and shared with damage information across other natural hazards.

Therefore, best practices codified by the International Standards Organisation (ISO) and the Open Geospatial Consortium (OGC) were adopted. Further information is available within Osuchowski and Atkinson (2008) and Atkinson et al. (2007).

Process

In order to map content from each database provider to the common schema, we needed to develop a series of rules or commands for the translation of data. This proved to be difficult due to the large number of free text descriptions in the databases. The entire contents of a free-text description needed to be mapped to a single term or number of terms in the interface. The use of free text fields also meant similar information was described differently within a single database. Consider for example: debris flow, debris-flow, debris/earth flow, or complex debris flow-earth slide. Each separate instance needed to be manually mapped to the common schema.

While a 'bottom up approach' enables data to be produced to a nationally consistent format from existing data, a 'top down' approach that encourages the use of standards in the development of new databases would provide greater functionality and also allow direct mapping from new database providers to the interface in future.

The result

Successful implementation of the methodology is demonstrated in connecting three physically separate and unique landslide databases via the web (www.ga.gov.au/landslide).

The most important advantage of adopting such an approach is the increased volume of information it facilitates. The database now has over 3630 entries detailing landslides and sets of landslides throughout Australia. Over 2074 landslides are being reported from MRT, over 1000 are reported from GA and 402 are reported from UoW.

Time and resource constraints dictated the level of functionality enabled as part of the pilot. The current LDIP is a demonstrator and further work is still needed to achieve a 'stable' system. These are further described in Osuchowski & Atkinson (2008). Examples include:

- a governance framework is required to manage changes to vocabularies. If new free-text descriptions are developed by custodians, the interface cannot map to this data;
- rules are also required to specify how the system behaves. For example if a connection to one of the three databases is temporarily unavailable, the search is aborted. Rules can specify the return of all data available, with a message indicating which database is unavailable; and
- performance optimisation of the application is needed as it can presently take up to one minute to execute a search.

Current benefits include:

- the system collates and characterises information from different sources in real time, providing an automatically updated single point of access to landslide information. New information is immediately available online. There is no need to wait for manual updates;
- data is presented consistently to enable the comparison and aggregation of data across databases;
- users are able to simultaneously search and query remote databases regardless of where they are hosted or differences in format, providing greater availability, accessibility and discoverability of data;
- detailed information can be accessed for specific requirements or generic information can be aggregated for strategic purposes. Drill-down functionality means different users can access the level of information they require from the same, single information-rich source;
- the need to locate, access and interrogate isolated databases or to separately identify and contact a number of individuals when information is needed is removed;
- the system provides an ability to export data in a range of formats, such as kml (Google Earth) or display results as reports, tables, maps and potentially as graphs and statistics;
- users can access multi-media such as photographs, videos, published papers, articles etc.;
- database custodians have greater flexibility and functionality in searching for their own data, and in comparing landslides occurring under similar conditions in other parts of the country for example. Custodians also select which fields of data they would like to share;
- databases can be connected to the interface whether or not they are available in an online capacity. For example, MRT and UoW do not have their landslide databases available separately online; and
- there is no limit to the number of landslide databases that can be linked into the virtual database since the interface neither stores or records data.

Discussion

At a minimum this demonstrator initiative provides Australia with a framework for a centralised national landslide inventory, which with further work could connect other available landslide databases in Australia. However, there is also considerable capacity for this approach to provide State Governments with a simple way to compile and maintain their own state-wide databases.

Interoperability is becoming increasingly relevant to federal government decision makers and research groups, all of whom need to access data and information across Australia through one system. This is especially the case in the research and management of natural hazards in Australia.

Implementing RC2 effectively and sustainably is a challenging task, but it is possible. The methodology of the LDIP has the capacity to be applied across to other hazards, such as flood, earthquake, tropical cyclone and bushfire. This would require a greater policy and governance framework, balanced with a greater technical capacity. The coordinated development of common vocabularies targeting requirements across all user groups would also be needed, but the benefits would be significant. Land-use planners, emergency managers, town planners, policy officers and researchers would be able to:

- access up-to-date information;
- access the same source data;
- share and compare methodologies;
- compare and contrast data within and between hazards; and
- engage in greater discussion on how to better reduce the risk to Australian communities from natural hazards.

Conclusions

The interoperable approach described here establishes a platform to support improved risk assessments and informed mitigation decisions through its ability to collate and characterise large volumes of information. In using a common data modelling methodology, the landslide domain model provides the capacity to extend the approach across other natural hazard databases and integrate data from other domains, leading to gains by all levels of government as well as academia and insurance organisations.

It is impractical and expensive for a single agency to maintain an up-to-date central database by collating and physically integrating data from different sources. An interoperable approach ensures that full value is made of available information, and that responsibility for collecting and maintaining this data is shared across all agencies. Specific-purpose data can not only continue to serve the needs of individual database custodians, but can also now serve a broader need. By sharing and exchanging data more efficiently we can also build more effectively on previous knowledge and reduce duplications in effort.

Such a system establishes the foundation for a very powerful and coordinated information resource in Australia and provides a suitable basis for greater investment in data collection, facilitating a sound knowledge base on natural disasters and disaster risk reduction.

References

Australian Geomechanics Society (2007) Guideline for landslide susceptibility, hazard and risk zoning for land use planning. Australian Geomechanics, 42 (1): 13-36.

Atkinson, R. Millard, K. and Arctur, D. (2007) Standards based approaches for cross domain data integration. International Journal of Spatial Data Infrastructure Research, 2: 74-89.

Council of Australian Governments (2004) Natural Disasters in Australia: Reforming mitigation, relief and recovery arrangements. A report to the Council of Australian Governments by a high-level officials group, August 2002. Department of Transport and Regional Services. Canberra.

Cox, S. and Richard, S. (2005) A formal model for the geologic time scale and global stratotype section and point, compatible with geospatial information transfer standards. Geosphere. 1(3): 119-137.

INSPIRE (www.ec-gis.org/inspire/) accessed 24 July 2008.

Osuchowski, M. and Atkinson, R. (2008) Connecting diverse landslide inventories for improved information in Australia. Proceedings of the 1st World Landslide Forum, Tokyo, Japan, 18–21 November 2008.

Popescu, M. E. (1994) A suggested method for reporting landslide causes. Bulletin of Engineering Geology and the Environment, 50 (1): 71-74.

Acknowledgements

The author would like to acknowledge the technical contributions of Rob Atkinson, Simon Cox and Steve Irwin in this pilot project and also those from GA's Information Services Branch.

Acknowledgements extend to project partners Colin Mazengarb (Mineral Resources Tasmania) and Phil Flentje (University of Wollongong), in addition to those who contributed to the development of the application schema. The 'virtual' Australian Landslide Database was also partly made possible through funding provided by Emergency Management Australia.

Finally, the author would like to thank Ole Nielson, Stuart Ross and Trevor Dhu who reviewed this paper.

About the author

Monica Osuchowski joined Geoscience Australia in 2004 after completing an Honours degree in Environmental Geology at the University of Tasmania. Monica is presently leading the landslide research program at GA.