Developing an Integrated and Interdisciplinary Information Platform for Volcanic Hazard and Risk Data

Nayembil, M.[1], Loughlin, S.[2], Duncan, M.[2] and the GVM Network

[1]British Geological Survey (a Natural and Environment Research Council Institute), Environmental Science Centre, Nicker Hill, Keyworth, Nottingham, UK, NG12 5GG (mln@bgs.ac.uk)

[2]British Geological Survey (a Natural and Environment Research Council Institute), Lyell Centre, Research Avenue South, Edinburgh, EH14 4AP (md@bgs.ac.uk)

In recent years, the volcanology community has begun to acknowledge the need to better organise and integrate their data to enable their research; however a key issue to resolve is how to integrate databases at a volcano or national scale in such a way that we can expand them to a regional and ultimately global scale and, through various tools, be able to share the data for effective research. By presenting examples of recent collaborations, we demonstrate the need to adopt common methodologies for data collation, integration and common vocabularies for QA procedures. Through the GVM network, we are moving towards an integrated and interdisciplinary information platform that brings together the multitude of volcanic datasets with data discovery, access and download tools for scientists. The ultimate goal is to enable linkages with other research data systems by providing programmable interfaces to extend data discoverability to regional and global scales for broader scientific research.

Keywords: volcanoes, databases, standardisation, vocabularies, relational, platform, interdisciplinary

In recent years, the volcanology community has begun to acknowledge the need to better organise and integrate their data to enable their research. A key issue to resolve is how to integrate databases at a volcano or national scale in such a way that we can expand them to a regional and ultimately global scale and, through various tools, be able to share the data for effective research. The answer lies in seeking a convergence on methodologies for data standardisation, harmonisation, discovery, availability and use. We demonstrate that it is only through collective cooperation that this can be achieved.

In order to understand and anticipate volcanic hazards and the related risk, we require extensive and divergent data. Volcanic data can include, amongst others, real-time and/or time-series data for monitoring the activity of volcanoes; hazard data related to the multiple hazards produced by volcanoes; historical data on previous eruptions, their magnitudes and impacts; as well as exposure and vulnerability data required to estimate risk. A series of challenges must therefore be overcome in order to enhance the management and integration of volcanic data. The challenges include: (1) the fact that data are disaggregated; (2) lack of data standards and standard vocabularies; (3) lack of metadata; and (4) data inaccessibility, as they often sit within personal data stores or organisations rather than in accessible information hubs.

Using the example of the Global Volcano Model (GVM) network, we provide case studies of how we have gone some way to address these challenges. The Global Volcano Model (GVM) network is a public–private network comprising local, national, regional and international organizations and institutions concerned with identifying and reducing aspects of volcanic risk and building the resilience of society to volcanic eruptions. An objective of GVM is to develop an accessible and sustainable information platform on volcanic hazards and risk. Through GVM, we have developed a spatially-enabled database with a web tool on global eruptions of magnitude 4 or greater, reaching back to the start of the Quaternary (LaMeve[1]). We have incorporated a methodology for data collation and standardisation, developed standard vocabularies and implemented an extensible data model to capture new data types. The web tools, in addition to the data discovery, access and download functions will be extended to provide a web service for direct harvesting by other systems. GVM's work, with its data standards, relational databases, web APIs (services) and the general ontological approach, will enable European research infrastructure projects and others to directly harvest volcanic hazards and risk data for the broader scientific community. The structure of the GVM platform is evolving and it currently takes the form of a series of connected relational databases addressing multiple volcanic hazards and impacts. The suite of tools so far developed and maintained by the community are diverse and mainly hazards modelling tools and platforms at local to national level. GVM intends to enhance current hazards and risk modelling ability with a strong focus on international collaboration and cooperation.

The ultimate goal is to develop an interdisciplinary volcanic hazards and risk data platform with standardised and linked databases supported by community/national/globally defined controlled vocabularies with a web portal for data discovery and access. The platform will also incorporate a series of tools: API's and web services for programmatic access to the hosted data by e-research infrastructures (e.g. EPOS) and by other scientific data discovery and access platforms benefitting the volcanology community, amongst others.

Work towards this goal is complicated by issues within volcanology itself; for instance a lack of basic data in many countries even to establish with some confidence the eruption histories of volcanoes (their frequency, magnitude, intensity, and eruption characteristics), as well as the fact that many active volcanoes are unmonitored or have minimal monitoring, so the ability to anticipate future events (forecasting and short-term risk analysis) on the basis of real-time monitoring is also compromised. As the demand for open-access data increases, ensuring databases and research outputs reflect the uncertainty and caveats inherent within volcanic data is critical.

It is crucial to enhance the interdisciplinary communications between scientists, data architects and end users in order to convey the aforementioned volcanic science and information science complexities and to enable appropriate data use.

[1]See: http://www.bgs.ac.uk/vogripa/view/controller.cfc?method=lameve