



Metadata in geomagnetism: experiences of the World Data Centre for Geomagnetism, Edinburgh

Sarah Reay (sjr@bgs.ac.uk) Ewan Dawson, Susan Macmillan, Simon Flower, Thomas Humphries and Ellen Clarke
British Geological Survey, West Mains Road, Edinburgh EH9 3LA, United Kingdom

1. Introduction

Geomagnetic observatory data date back more than 150 years. Traditionally these data were stored with relevant metadata in the form of annual yearbooks. Metadata are information about data; they are the "who, what, when, where, why and how" of a data set. They describe the content, quality, originator, and other characteristics of a data set that help users understand the nature of that data set and how to use it.

The type of information required of an observatory metadata record includes, but is not limited to:

- **Contact Information** - The name of the responsible institute that produces the data, including addresses and personal contacts.
- **Data Description** - A description of the type of data, including a narrative summary of the nature of the data and its possible applications, and a list of entities and attributes relevant to the data.
- **Station Description** - A description of the station in question, including coordinates, elevation, and possibly photographs and maps.
- **Instrumentation** - A description of the types of instruments in use at the observatory, and the nature of the data that are retrieved from them.
- **Data Processing** - A description of the processes and methodology used to process the data from instrument recordings to the final definitive values.
- **Data Quality** - A general assessment of the quality of the data set considering completeness, accuracy, quality control methods and so on.
- **Data Distribution** - A description of how and from where the data may be acquired.

Since the advent of digital recordings in the 1960s metadata have become increasingly separated from the actual data. Digital data are stored in various formats (commonly IAGA-2002, WDC format, INTERMAGNET binary) which have scope for only very limited metadata information. Additionally many organisations have stopped producing yearbooks so there is currently no place where much of these metadata are formally recorded.

2. Why does this matter?

It is important that current and future generations have sufficient information about the data to enable them to independently understand and use them. Currently there is little information on the provenance of a digital data set leading to inconsistencies where a data set may have been modified or corrected at one data outlet but not at others. Without sufficient metadata it soon becomes difficult or impossible to determine what transformations or corrections have been applied. Better metadata would provide a clear record of any work and provide quality assurance to researchers using these data.

If metadata were stored in a common metadata standard this would have further benefit allowing better data discoverability and interoperability with e-science initiatives being established such as the Global Earth Observation System of Systems (GEOSS), the European Plate Observing System (EPOS) and the World Data System (WDS).

3. Metadata standards

With WDCs Boulder and Kyoto we have discussed and examined the idea of establishing a metadata standard for geomagnetic data over the past few years. We have encountered many issues with this.

The principle difficulty is in applying a standard to time-series data. Common well-established metadata standards such as ISO-19115 and FGDC are based on spatial data which do not change with time. With geomagnetic data however the metadata changes with time; most commonly as a result of a change in instrumentation or data processing and therefore data quality. More complex changes such as pillar changes would also have to be accounted for somehow.

One other issue is complexity; a complete metadata record is long with many fields to be completed and the language of metadata is not always readily understandable to the non-expert.

Encouraging institutes operating observatories to adopt and populate standard metadata record is challenging.

Association: IAGA & IASPEI
Symposium: J-A05
Poster: 5780 (Sat 2nd July)

IUGG Melbourne, Australia
28th June - 7th July 2011

4. What are we doing?

Due to the difficulty in progressing with metadata standards we have focussed our attention simply on the acquisition of metadata in whatever form it takes. We decided to first collect the information and then import this into a standard, when one is established.

Each year the World Data Centre (WDC) for Geomagnetism, Edinburgh issue a 'call for data' requesting the latest data from operating observatories. In the past two years we have also added a request for simple metadata.

We have been collecting and/or checking an observatory's name, IAGA code, dates of operation, location, contact information, instruments used, INTERMAGNET status etc. We store this information in a Microsoft Access database and make the information available from our 'Data Catalogue' website (www.wdc.bgs.ac.uk).

We have also been asking for other sources of metadata. We have put links to an observatory's website, if available, as a potential source of further metadata. We have requested copies or links to yearbooks if available. Yearbooks remain the most complete metadata record currently available and it is good to associate these with the digital data sets. We have also developed a facility on Google Maps so users can view an observatory's metadata or search for observatories by location.

In addition we have begun to address the issue of data discrepancies and unknown data processing steps. We have a new field which we have begun to populate with any information we have on a data set. This is a free form text box to note any known issues. As a start publicising these issues will help users to better know the data they are working with, will inform observatory operators that there may be an issues and give the WDC a place to note any modification they have made to a data set.

We welcome feedback from users of WDC data to help us add to this record.

World Data Centre for Geomagnetism (Edinburgh)

Observatory Details

LAGA Code	BEL
Name	Belk
Opened	1960-01-01
Closed	
Latitude	51.837°N
Longitude	20.792°E
Altitude	180.0m
Country	Poland
Website	http://www.wdc.bgs.ac.uk/observatory.html
INTERMAGNET Member	1993 - present
Notes	

Contact Details

Name	Jan Reida
Address	Central Geophysical Observatory 05-622 Belk POLAND
E-mail	janr@bgs.ac.uk
Fax	+48 48 6610940

Name	Mariusz Neska
Address	Central Geophysical Observatory 05-622 Belk POLAND
E-mail	neska@bgs.ac.uk
Fax	+48 48 6610940

Yearbooks

2008	http://www.wdc.bgs.ac.uk/yearbooks/yearbook_2008.pdf
2007	http://www.wdc.bgs.ac.uk/yearbooks/yearbook_2007.pdf
2006	http://www.wdc.bgs.ac.uk/yearbooks/yearbook_2006.pdf

Instrumentation

Instrument	Type	In Use From	In Use To
FMP-8 proton precession magnetometer (Institute of Geophysics PAS)	Scalar		
RSM Bobrov-type quartz magnetometers (Institute of Geophysics PAS)	Vector (OXYZ)		
DL-Builder magnetometer (type ELSEC 810)	Absolute		

Free-form text space for noting any known issues with the data e.g. data quality, discrepancies or any corrections applied.

Current instruments used by an observatory. There are also fields to note the dates these were in use so this could be used to capture a time-varying aspect of an observatory's metadata.

5. What next?

We aim to continue populating these records. In the last 'call for data' we began contacting operating observatories directly to try and gather some missing information. We will populate our metadata with any known data issues imminently. We are also gathering metadata from INTERMAGNET DVDs. We wish to encourage the production of yearbooks and would welcome submission of these. We wish to consider if the field of data mining may be used to extract metadata direct from yearbooks. When records are well populated we would like to work back in time, especially with instrumentation records.

We also encourage discussion on metadata standards and the recent involvement of INTERMAGNET to help solve the many issues surrounding this topic and would welcome advice from other science areas that may have already tackled a similar issue.

References

Fischman D, Denig W F and Herzog D (2009) A Proposed Metadata Implementation for Magnetic Observatories. Proceedings of the XIIIth IAGA Workshop on Geomagnetic Observatory Instruments, Data Acquisition, and Processing: U.S. Geological Survey Open-File Report 2009-1226, 82-85 pp

Reay S J, Herzog D C, Alex S, Kharin E, McLean S, Nosé M and Sergeyeva N (2011) Magnetic Observatory Data and Metadata: Types and Availability. In: Mandea M, Korte M (eds) Geomagnetic Observations and Models, IAGA Special Sopron Book Series 5, DOI 10.1007/978-90-481-9858-0_7 pp149-181