



Hands On
Data from the World Data
Centre for Geomagnetism

Lee Bargatze
Todd King



The World Data Center for Geomagnetism

- **The Seismology and Geomagnetism programme of the British Geological Survey hosts the World Data Centre for Geomagnetism (Edinburgh) as part of the World Data Centre (WDC) system established by the International Council of Scientific Unions (ICSU). The WDC has a comprehensive set of digital geomagnetic data as well as indices of geomagnetic activity supplied from a worldwide network of magnetic observatories. The data and services at the WDC are available for scientific use without restrictions.**
- <http://www.wdc.bgs.ac.uk/catalog/intro.html>



Data Access and Storage

- Available from an FTP server.
 - WDC-G provides an "FTP helper" to aid in navigating the FTP site.
- Hourly files
 - One WDC file for each station and year.
- Minute Files
 - Data is stored in ZIP files
 - One for each station and year.
 - ZIP file contains 12 files one for each month.



WDC Data Format

- Data is striped
 - multiple values for a single parameter packed into a record.
- Slightly different format for the 1-hr and 1-min cadence files.

1-hr cadence file format:

COLUMNS FORMAT DESCRIPTION

1-3	A3	Observatory 3-letter code
4-5	I2	Year. Last 2 digits, 1996 = 96. See also columns 15-16.
6-7	I2	Month (01-12)
8	A1	Element (D,I,H,X,Y,Z, or F)
9-10	I2	Day of month (01-31)
11-12	A2	Blanks
13	A1	Arbitrary
14	A1	Arbitrary
15-16	I2	Century digits
		Year = 1889, Century digits = 18
		Year = 1996, Century digits = 19
		Year = 2014, Century digits = 20
17-20	I4	Tabular base.
		The values are in degrees for D and I, and in hundreds of nanoTeslas for the intensity elements.
21-116	24I4	24 4-digit hourly mean values for the day.
		The values are in tenth-minutes for D and I, and in nanoTeslas for the intensity elements.
		The first hourly mean value represents the mean value between 00:00 UT and 01:00 UT, ..., the 24th value represents the mean between 23:00 UT and 24:00 UT.
		A missing value is identified by 9999.
117-120	I4	Daily Mean.
		If any of the hourly mean values for the day are missing 9999 will appear as the daily mean.
121-122		Record end marker.
		Two chars 'cr'= 13 and 'nl'= 10.

```

*== OLD FORMAT =====*
| Column 15: INTERNATIONAL QUIET |
| or DISTURBED DAYS Q=1, D=2 |
| Column 16: Blank for data since 1900, |
| 8 for data before. |
*=====*
```



Challenges Creating SPASE Descriptions

- The AccessURL in NumericalData is to a folder not a file.
 - A user may be able to figure out what to do, but a service needs more information.
 - What is the <Format> (and <Encoding>) for a folder?
- Granules are really files within zip files.
 - How to represent this as an URL?
 - Granule <Format> is not the same as declared in NumericalData.
- Non-tabular data.
 - Simple format, requires interpolation for time values.
 - Parameter keys exist, so mapping is possible.
 - No parameter key for time.
- Temporal Parameters
 - How to describe format of time value?



Abisko Ground Station Numerical Data

```
<NumericalData>
  <ResourceID>spase://VMO/NumericalData/WDC/Abisko/Magneto
  <ResourceHeader>
    <ResourceName>Abisko Fluxgate Magnetometer Data</Resour
    <AlternateName>None</AlternateName>
    <ReleaseDate>2007-06-22T22:45:17.000</ReleaseDate>
    <ExpirationDate>2099-12-31T23:59:59.999</ExpirationDate>
    <Description>Abisko Magnetometer PT1M data; Station Id: ABK
    <Acknowledgement>Geological Survey of Sweden (SGU)</Ackn
    ...
  <InformationURL>
    <Name>WDC for Geomagnetism, Edinburgh; Geomagnetic Da
    <URL>http://www.wdc.bgs.ac.uk/catalog/intro.html</URL>
    <Description>Download web site for magnetometer data from th
  </InformationURL>
  <AssociationID>spase://VMO/Instrument/WDC/Abisko/Magnetom
  <PriorID>None</PriorID>
</ResourceHeader>
<AccessInformation>
  <RepositoryID>spase://VMO/Repository/WDC</RepositoryID>
  <Availability>Online</Availability>
  <AccessRights>Open</AccessRights>
  <AccessURL>
    <Name>FTP Site of the WDC at the British Geological Survey;
    <URL>ftp://ftp.nmh.ac.uk/wdc/obsdata/1minval/</URL>
    <Description>FTP site for downloading ground magnetogram d
  </AccessURL>
  <Format>Text</Format>
  <Encoding>ZIP</Encoding>
  ...
</NumericalData>
```

```
<PhysicalParameter>
  <Name>Abisko PT1M Magnetic Field Vector</Name>
  <Description>Abisko PT1M Magnetic Field Vector</Description>
  <Cadence>PT1M</Cadence>
  <Units>nT</Units>
  ...
  <Structure>
    <StructureType>Vector</StructureType>
    <Size>3</Size>
    <Description>Magnetic Field Vector Data</Description>
    <Element>
      <Name>Bx</Name>
      <Component>X</Component>
      <Index>1</Index>
      <ParameterKey>X</ParameterKey>
    </Element>
    ...
  </Structure>
  <Measured>
    <Field>
      <FieldQualifier>Vector</FieldQualifier>
      <FieldQuantity>Magnetic</FieldQuantity>
    </Field>
  </Measured>
</PhysicalParameter>
<PhysicalParameter>
  <Name>Time</Name>
  <Description>Timestamp for sample</Description>
  <Support>Temporal</Support>
</PhysicalParameter>
</NumericalData>
```



Abisko Ground Station Granule

```
<Granule>
  <ResourceID>spase://VMO/Granule/WDC/Abisko/Magnetometer/PT1M/1979/Jan</ResourceID>
  <ReleaseDate>2007-06-22T22:45:17.000</ReleaseDate>
  <ExpirationDate>2099-12-31T23:59:59.999</ExpirationDate>
  <ParentID>spase://VMO/NumericalData/WDC/Abisko/Magnetometer/PT1M</ParentID>
  <PriorID>None</PriorID>
  <URL>ftp://ftp.nmh.ac.uk/wdc/obsdata/1minval/1979/abk1979m.zip</URL>
  <StartDate>1979-01-01T00:00:00.000</StartDate>
  <StopDate>1979-12-31T23:59:00.000</StopDate>
  <Checksum>
    <HashValue>977b62e310f9fc8f055a7270f07e4a65</HashValue>
    <HashFunction>MD5</HashFunction>
  </Checksum>
  <DataExtent>
    <Bytes>1193376</Bytes>
    <Units>Bytes</Units>
    <Per>P1M</Per>
  </DataExtent>
</Granule>
```



Suggested Solutions

- Services, services, services...
 - Standard services could be used everywhere.
 - WDC reader (subset, extract) – available in MANGO
 - FTP site extractor (Candey?)
 - Zip file extractor (URL to retrieve a file from a ZIP)

Critical Changes

- Move <Format> and <Encoding> under <AccessURL>
- Change <URL> in <Granule> to <AccessURL>

Very Helpful Changes

- Include Accessor information in SPASE description.
 - <AccessURL>
 - <AccessorID>...

Note: <AccessorID> refers to a Service resource.



Conclusions

- SPASE descriptions are rich enough to describe a very challenging set of data.
- Minor refinements could enable more value added services.