

Data Wrangling

A high Time-Consuming task

Examples (from BGS)

Rosa Filgueira Vicente
12,06,2017

British Geological Survey

Background of BGS

World-leading geological survey.

- Public-good science for government, and research to understand earth and environmental processes
- Geo-scientific data providers

Informatics - Science Directorate (3 sections)

- Data Science
- National Geological Repository
- Information Systems

Huge breadth of data and information

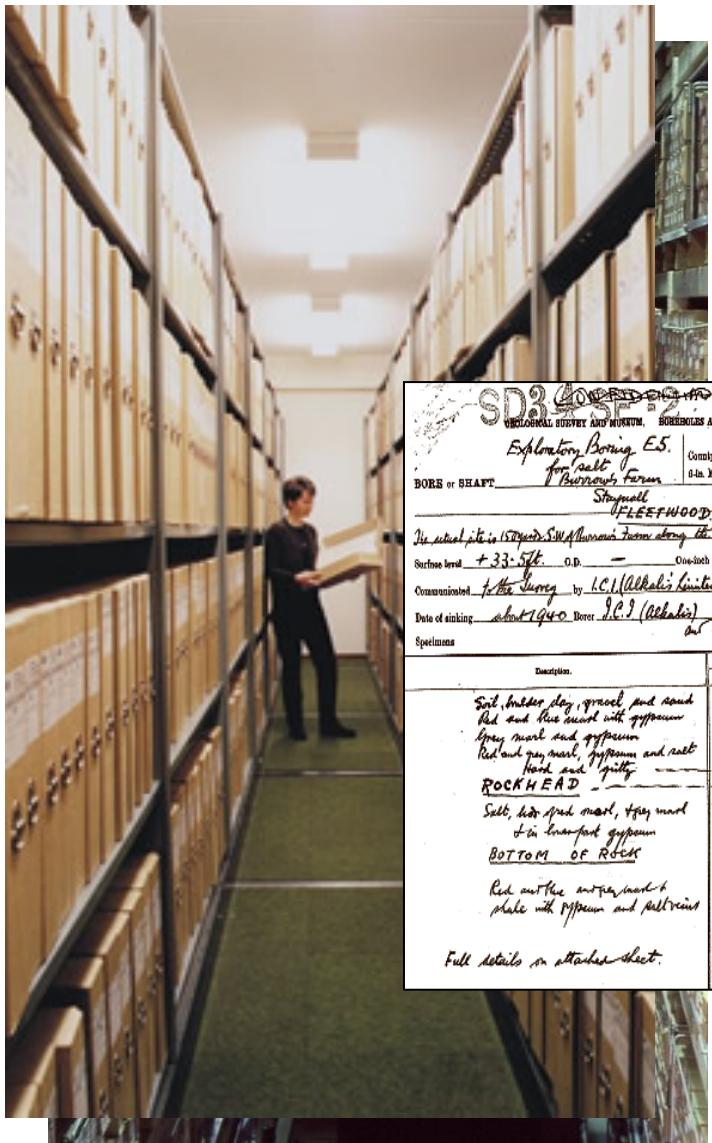
- Digital: 200TB and growing by the month (SAN)
- Records: 17.5 linear km's of paper records, maps, notebooks and site investigation reports

Background of BGS

The screenshot shows the homepage of the British Geological Survey (BGS) website at www.bgs.ac.uk. The page features a top navigation bar with links for Home, Data, Research, Services, Contacts, Discovering geology, News and events, Hosted sites, and a search icon. Below the navigation is a banner for 'Opengeoscience' with a list of services like View maps, Apps, Map data downloads, Photos and images, Publications, Scanned records, Data collections, Software, and Web services. To the right of this is a section for 'Our products' listing Energy, Environmental chemistry, Geology, Geophysics, Ground conditions, Groundwater, Hazards, Marine, and Minerals. Further right is a 'Top data services' section with links to Discovery metadata, Earthwise, GeoIndex, Geology of Britain viewer, iGeology, Lexicon, Licensing, Maps portal, and Search our data. On the far right, there are sections for the National Geoscience Data Centre (listing Cited data, Deposit data with NGDC, Geology data management, Groundwater data, information and services, Marine data access, and National Hydrocarbons Data Archive), the National Geological Repository (listing What's in the NGR?, Facilities, Materials collection, Records & archives, and Users), and a 'News' sidebar with recent articles about Michael Watts elected BGSN President, Depositing and accessing AGS and geotechnical data, Annual science review 2016-17, Limestone landscapes, Climate through time online, and New research using hydrogen aims to take the carbon out of heating. The main content area includes a 'Downloads' sidebar with links to Browse, Search, and Software disclaimer, and a large central image showing a geological map titled 'Climate through time online'.

www.bgs.ac.uk/data/home.html?src=topNav

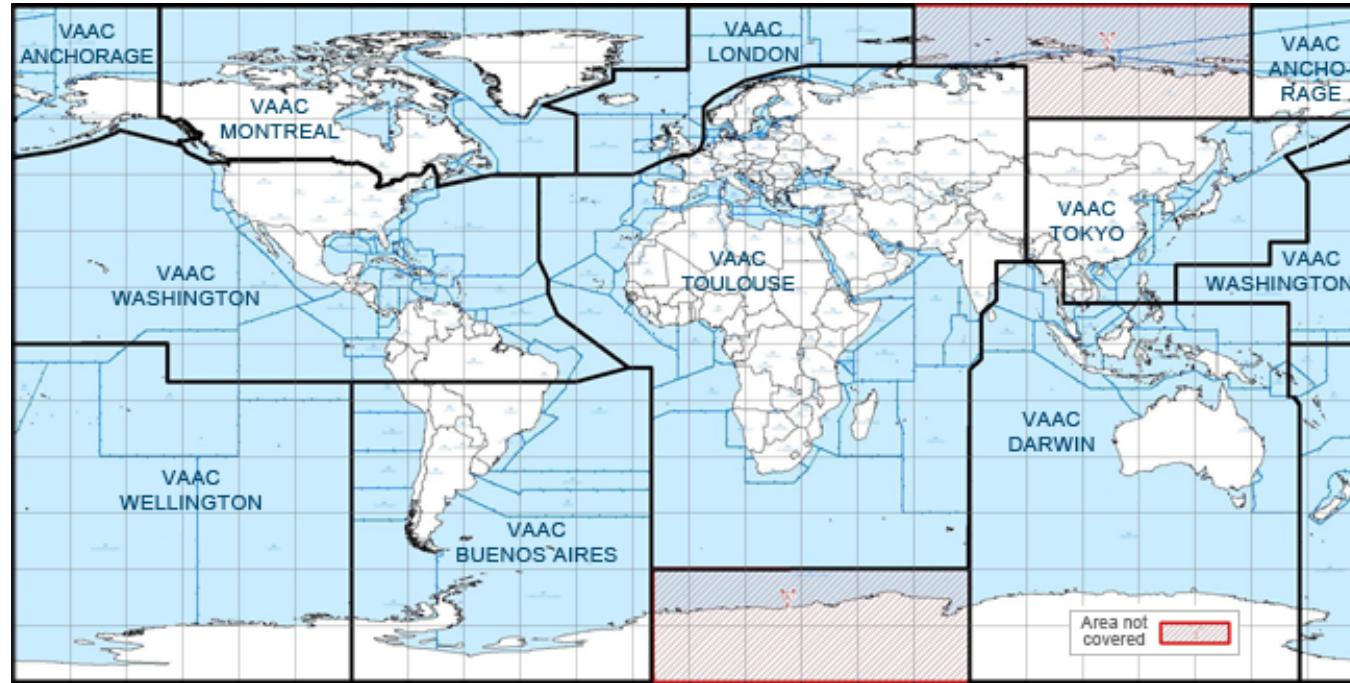
Breadth of Data and Information



Data Wrangling- Examples

Volcanology – Automatic information retrieval
Seismic data- Real-time analysis/monitoring

Volcanology: Volcanic Ash Advisories



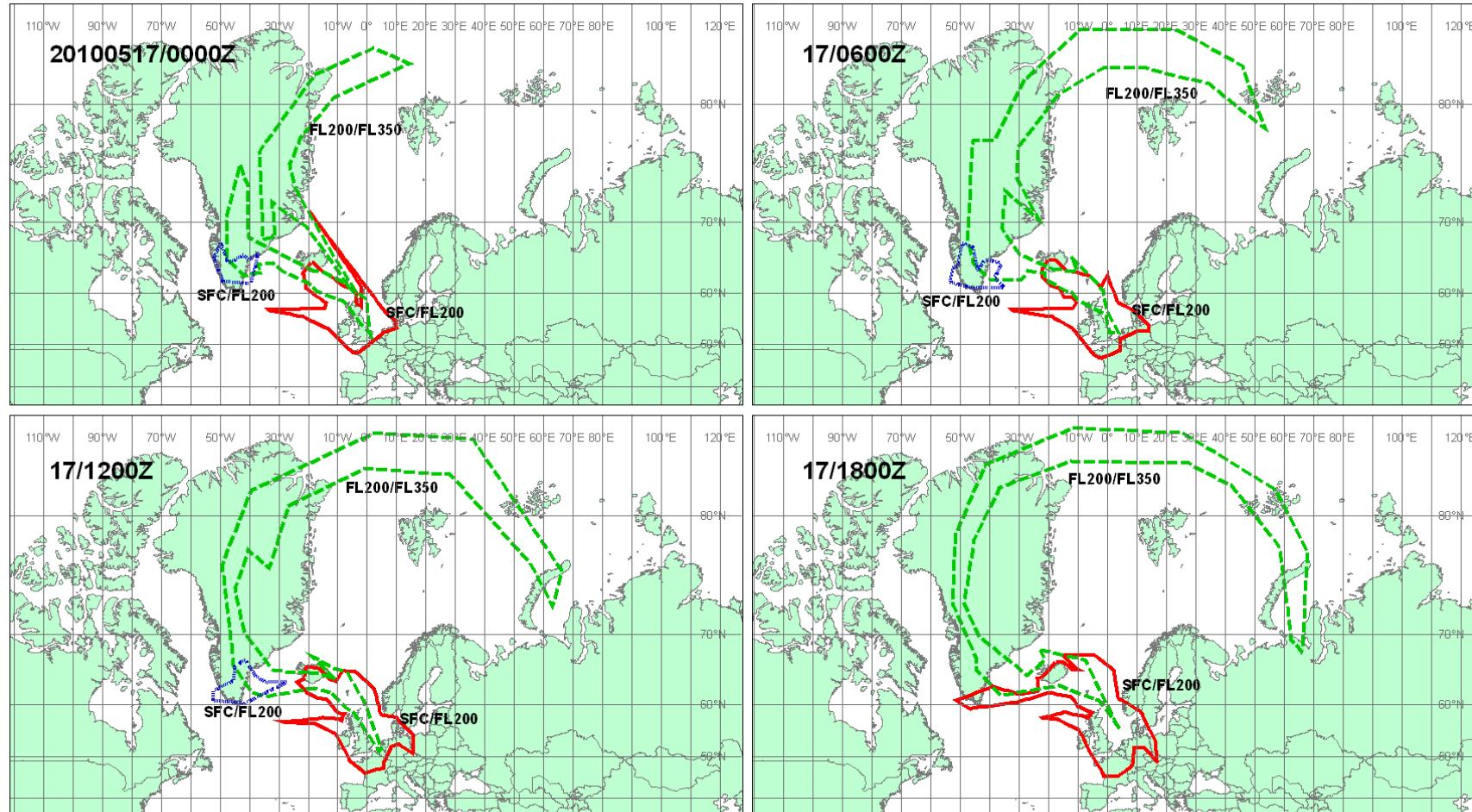
Information Sources:
Pilot reports
Volcano observatories
Satellite imagery

Volcanic Ash Advisory Centres:
Initiate dispersal simulations
Validate simulation results
Provide volcanic ash advisory

Aviation Sector:
Use information to inform flight paths

Volcanic Ash Advisories

Forecast ash dispersion



VA ADVISORY
DTG: 20100517/0000Z
VAAC: LONDON
VOLCANO:
EYJAFJALLAJOKULL
PSN: N6338 W01937
AREA: ICELAND

SUMMIT ELEV: 1666M
ADVISORY NR: 2010/126
INFO SOURCE: ICELAND MET OFFICE
AVIATION COLOUR CODE: RED
ERUPTION DETAILS: ERUPTION CONTINUING
TO AN ESTIMATED HEIGHT OF FL230 TO FL260,
OCNL FL290.

RMK: A SEPARATE AREA OF SFC TO FL200 OVER GREENLAND IS
MARKED IN BLUE ON GRAPHICS.
NXT ADVISORY: 20100517/0600Z

text providing key information on
the volcano, eruption and source of inf.

Volcanic Ash Advisories – What we can do with them?

ESP database

NUMBER	NAME	LOCATION	STATUS	LATITUDE	NS	VF	LONGITUDE	EW	ELEV	TYPE	TIME FRAME	ERUPTION TYPE
0201-041	Dallol	Ethiopia	Historical	14.242	N		40.30	E	-48	Explosion craters	D2	S0
0201-04=	Alid	Ethiopia	Holocene	14.88	N		39.92	E	904	Stratovolcano	U	S0
0201-05=	Gada Ale	Ethiopia	Holocene	13.975	N		40.408	E	287	Stratovolcano	U	M0
0201-06=	Alu	Ethiopia	Holocene	13.825	N		40.508	E	429	Fissure vents	U	M0
0201-071	Borale Ale	Ethiopia	Holocene	13.725	N		40.60	E	668	Stratovolcano	U	M0
0201-07=	Dalaffilla	Ethiopia	Historical	13.792	N		40.55	E	613	Stratovolcano	D1	M0
0201-08=	Ert A Ale	Ethiopia	Historical	13.60	N		40.67	E	613	Shield volcano	D1	M1
0201-091	Hayli Gubbi	Ethiopia	Holocene	13.50	N		40.72	E	521	Shield volcano	U	M0
0201-09=	Ale Bagu	Ethiopia	Holocene	13.52	N		40.63	E	1031	Stratovolcano	U	M0
0201-101	Nabro	Ethiopia	Holocene?	13.37	N		41.70	E	2218	Stratovolcano	?	S0
0201-102	Mallahle	Ethiopia	Holocene?	13.27	N		41.65	E	1875	Stratovolcano	?	S0
0201-103	Sork Ale	Ethiopia	Holocene?	13.18	N		41.725	E	1611	Stratovolcano	?	M0
0201-104	Asavyo	Ethiopia	Holocene	13.07	N		41.60	E	1200	Shield volcano	U	M0
0201-105	Mat Ala	Ethiopia	Holocene	13.10	N		41.15	E	523	Shield volcano	U	M0
0201-106	Tat Ali	Ethiopia	Holocene	13.28	N		41.07	E	700	Shield volcano	U	M0
0201-107	Borawli	Ethiopia	Holocene	13.30	N		40.98	E	812	Stratovolcano	U	M0
0201-10=	Dubbi	Ethiopia	Historical	13.58	N		41.808	E	1625	Stratovolcano	D3	M0
0201-111	Ma Alalta	Ethiopia	Holocene	13.02	N		40.20	E	1815	Stratovolcano	U	S0
0201-112	Alayta	Ethiopia	Historical	12.88	N		40.57	E	1501	Shield volcano	D2	M0
0201-113	Dabbahu	Ethiopia	Historical	12.60	N		40.48	E	1442	Stratovolcano	D1	S0
0201-114	Dabbayra	Ethiopia	Holocene	12.38	N		40.07	E	1302	Shield volcano	U	M0
0201-115	Manda Hararo	Ethiopia	Historical	12.17	N		40.82	E	600	Shield volcanoes	D1	M0
0201-116	Groppi	Ethiopia	Holocene	11.73	N		40.25	E	930	Stratovolcano	U	S0
0201-11=	Afderà	Ethiopia	Holocene?	13.08	N		40.85	E	1295	Stratovolcano	?	S0

Volcanic Ash Advisories – Problems – Data Wrangling

- **Getting** the data from the 9 VAAC
- **Extract** the desired information (> 60000 records)
 - **Different data distributions:** FTP servers, websites, intranets etc.
 - **Different data organization:** Per year, Per location, Per volcano ...
 - **Different presentation of datasets :** Tables, Text files, emails, etc.
 - **Different “naming” and ‘temporal’ IDs.**
 - **Different order of values inside the files & file formats**
 - (Variations between each VAAC and sometimes within the same one)

VAAC	<i>Website with reports</i>	<i>Dates of reports</i>
Anchorage	http://vaac.arh.noaa.gov/list_vaas.php	Since 2008
Washington	http://www.ssd.noaa.gov/VAAC/messages.html	Since 2007
London	http://www.metoffice.gov.uk/aviation/vaac/vaacuk_vag.html	Since 2010
Wellington	http://www.metoffice.gov.uk/aviation/vaac/vaacuk_vag.html	Since 2001
Buenos Aires	http://www.smn.gov.ar/vaac/buenosaires/productos.php?lang=en&anio=2011	Since 2011
Darwin	ftp://ftp.bom.gov.au/anon/gen/vaac/	Since 1998
Toulouse	http://www.meteo.fr/vaac/eva.html	Since 2008
Tokyo	http://ds.data.jma.go.jp/svd/vaac/data/index.html	Since 2006
Montreal	http://weather.gc.ca/eer/vaac/index_e.html	Since 2009

Volcanology – Information to extract



Each advisory constitutes a text file containing the following information:

VA ADVISORY

DTG: 20150223/2302Z [Date and time issuance of advisory]
VAAC: WELLINGTON [Responsible VAAC]
VOLCANO: AMBRYM 257040 [Volcano name and no]
PSN: S1615 E16807 [Coordinates of volcano]
AREA: VANUATU [Volcano area]
SUMMIT ELEV: 1334M [Elevation of volcano summit]
ADVISORY NR: 2015/17 [Advisory number]
INFO SOURCE: PIREPS AND SATELLITE SO2 IMAGERY [Source of information]
AVIATION COLOUR CODE: UNKNOWN [Typically provided by local VO's]
ERUPTION DETAILS: NO OBSERVATIONS OF SIGNIFICANT ASH SINCE PIREP
AT 212100Z OF ASH TO 8000FT. [Information regarding information]
OBS VA DTG: 23/2300Z [Date and time of observed ash]
EST VA CLD: VA NOT IDENTIFIABLE FM SATELLITE [Estimated height (FL) and extent of cloud]
FCST VA CLD+6 HR: 24/0500Z SFC/FL600 NO VA EXP [Forecast of plume height & extent in 6 hrs]
FCST VA CLD+12 HR: 24/1100Z SFC/FL600 NO VA EXP [Forecast of plume height & extent in 12 hrs]
FCST VA CLD+18 HR: 24/1700Z SFC/FL600 NO VA EXP [Forecast of plume height & extent in 18 hrs]
RMK: SATELLITE IMAGERY SHOWS NO SIGN OF ASH. NO FURTHER ADVISORIES WILL BE ISSUED UNTIL CONFIRMED REPORTS RECEIVED. [Further info. on eruption and observations]
NXT ADVISORY: NO FURTHER ADVISORIES= [Time of next advisory]

Volcanology – 50 % Files have this format



IDD41290

VA ADVISORY

DTG: 20160106/0330Z

VAAC: DARWIN

VOLCANO: SINABUNG 261080

PSN: N0310 E09824

AREA: INDONESIA

SUMMIT ELEV: 2460M

ADVISORY NR: 2016/1

INFO SOURCE: GROUND REPORTS, HIMAWARI-8

AVIATION COLOUR CODE: RED

ERUPTION DETAILS: MINOR VA/STEAM EMISSION TO FL140

EXTENDING TO THE WEST

OBS VA DTG: 06/0330Z

OBS VA CLD: SFC/FL140 N0314 E09824 – N0310 E09810 – N0305

E09811 – N0304 E09817 – N0310 E09826 MOV SW 5KT

FCST VA CLD +6 HR: 06/0930Z SFC/FL140 N0314 E09824 – N0313

E09811 – N0304 E09754 – N0248 E09738 – N0235 E09737 – N0222

E09748 – N0237 E09809 – N0309 E09827

FCST VA CLD +12 HR: 06/1530Z NO VA EXP

FCST VA CLD +18 HR: 06/2130Z NO VA EXP

RMK: VA EXPECTED TO DISSIPATE WITHIN 12 HRS.

NXT ADVISORY: NO LATER THAN 20160106/0930Z

Copyright Commonwealth of Australia 2011, Bureau of Meteorology (ABN 92 637 533 532). Users of these web pages are deemed to have read and accepted the conditions described in the Copyright, Disclaimer, and Privacy statements

Volcanic – Other 50 % have different formats , and name conventions

```
&nbsp; &nbsp; &nbsp; Home  

&nbsp; &nbsp; &nbsp; Accueil  

FVXX01 LFPW 062058  

VOLCANIC ASH ADVISORY  

ISSUED: 20070406/2100Z  

VAAC: TOULOUSE  

VOLCANO: PITON DE LA FOURNAISE 0303-02  

LOCATION: S2113E05543  

AREA: REUNION ISLAND ON INDIAN OCEAN  

SUMMIT ELEVATION: 2631M  

ADVISORY NUMBER: 2007/03  

INFORMATION SOURCE: SATELLITE IMAGERY  

AVIATION COLOUR CODE:  

ERUPTION DETAILS: TONIGHT, AND SINCE THE CRATER COLAPSE HAS BEEN  

OBSERVED THIS  

AFTERNOON, IT SEEMS THAT THERE IS NO INCREASING ACTIVITY ;  

SATELLITE IMAGERY SHOWS A DURING SIMILAR PLUME THAN TODAY, PROBABLY  

CONTAINING SOME ASHES.  

OBS ASH DATE/TIME: 06/2030Z  

OBS ASH CLOUD: SFC/FL180 S2013 E05543 - S2120 E05610 - S2100 E05630  

-S2040 E05616  

- S2013 E05543  

FCST ASH CLOUD+6H: SFC/FL180 S2013 E05543 - S2100 E05600 - S2100  

E05600 -S2040 E05547  

- S2013 E05543  

FCST ASH CLOUD+12H: SFC/FL180 S2013 E05543 - S2100 E05600 - S2100  

E05600 -S2040 E05547  

- S2013 E05543  

FCST ASH CLOUD+18H: SFC/FL180 S2013 E05543 - S2100 E05600 - S2100  

E05600 -S2040 E05547  

- S2013 E05543  

NEXT ADVISORY: NEXT ADVISORY WILL BE ISSUED 20070407 AT 0300Z  

REMARKS: PLEASE CHECK STORMET FOR CURRENT WARNINGS
```

```
<?xml version="1.0"?>  

<volcanicash>  

<wmoheader>FVPS01 NZKL 050000</wmoheader>  

<body><![CDATA[  

VA ADVISORY  

DTG: 20110705/0000Z  

VAAC: WELLINGTON  

VOLCANO: CORDON CAULLE 1507-141  

PSN: S4031 W07212  

AREA: CHILE C  

SUMMIT ELEV: 1798M  

ADVISORY NR: 2011/128  

INFO SOURCE: DARWIN VAAC, MTSAT, HYSPLIT, MODIS  

ERUPTION DETAILS: ONGOING ERUPTION  

COLOUR CODE: NIL  

OBS VA DTG: 05/0000Z  

EST VA CLD: SFC/FL200 NO VA EXP FL200/400 S5700 E16000 - S5200 W15500 - S4145 W13500 -  

- S7030 W15745 - S7000 E16000 - S5700 E16000 FL400/600 NO VA EXP  

FCST VA CLD+6H: 05/0600Z SFC/FL200 NO VA EXP FL200/400 S5900 E16000 - S5300 W15400 - S4615  

5 W15600 - S7000 E16000 - S5900 E16000 FL400/600 NO VA EXP  

FCST VA CLD+12H: 05/1200Z SFC/FL200 NO VA EXP FL200/400 S4900 W12900 - S7200 W11115 - S7230  

0 W12900 FL400/600 NO VA EXP  

FCST VA CLD+18H: 05/1800Z SFC/FL200 NO VA EXP FL200/400 S6000 E16000 - S5100 W13045 - S7030  

0 E16000 FL400/600 NO VA EXP  

RMK: FOR VA W OF E16000 SEE DARWIN VAA FVAU01.  

NXT ADVISORY: NO LATER THAN 20110705/0600Z=
```

]]></body>

Just
FEW
examples

Subject:
VolcanoMessage
From:
DBNet Account <dbnet@ncosrv2.ncep.noaa.gov>
Date:
Thu, 31 May 2007 14:01:30 +0000 (GMT)
To:
W-Vaac@noaa.gov, xogm@afwa.af.mil


658
FVXX21 KNES 311357
VOLCANIC ASH ADVISORY
ISSUED: 20070531/1355Z VAAC: WASHINGTON

VOLCANO: COLIMA 1401-04
LOCATION: N1930 W10337 AREA: MEXICO
SUMMIT ELEVATION: 12631 FT (3850 M)
ADVISORY NUMBER: 2007/010
INFORMATION SOURCE: GOES-12. NAM WINDS.
ERUPTION DETAILS: BRIEF EMISSION AROUND 1100Z
OBS ASH DATE/TIME: 31/1332Z
OBS ASH CLOUD: SFC/FL170 10NM WIDE LINE BETWEEN
N1932 W10346 - N1902 W10359. ASH IS MOVING S 5-10
KNOTS.
FCST ASH CLOUD +6H: 31/1930Z SFC/FL170 NO ASH EXP.
FCST ASH CLOUD +12H: 01/0130Z NO ASH EXP.
FCST ASH CLOUD +18H: 01/0730Z NO ASH EXP.

REMARKS: IN THE MORNING VISIBLE IMAGERY A FAINT

FVFF01 R1TD 140149 2013045 0151
VA ADVISORY
DTG: 20130214/0149Z
VAAC: TOKYO
VOLCANO: SAKURAJIMA 0802-08
PSN: N3135E13040
AREA: JAPAN
SUMMIT ELEV: 1060M
ADVISORY NR: 2013/158
INFO SOURCE: JMA
AVIATION COLOUR CODE: NIL
ERUPTION DETAILS: EXPLODED AT 20130214/0139Z OVER FL040 EXTD NE
OBS VA DTG: 14/0115Z
EST VA CLD: VA NOT IDENTIFIABLE FM SATELLITE DATA WIND FL180 270/58KT
FCST VA CLD +6 HR: NO VA EXP
FCST VA CLD +12 HR: NO VA EXP
FCST VA CLD +18 HR: NO VA EXP
RMK: WE WILL ISSUE FURTHER ADVISORY IF VA IS DETECTED IN SATELLITE
IM
AGERY.
NXT ADVISORY: NO FURTHER ADVISORIES=

Volcanic Ash Advisories- Solution

Solution: Created a HUGE parser script (Python) that deals with all the exceptions – Time needed for created this script- 2 Months and several (small) improvements during the next 6 months.

Key for success- Workflow of communication:

- Identify WITH the domain expert the values to be captured (which data are essential, optional, etc.)
- Find the different versions for each of these values: naming, positions in the file, if can or cannot to appear at all.
- Create a first version and iterate from it
- Allow the domain expert to identify more problems and solve them as soon as possible

Volcanic Ash Advisories- Solution

Data-pipeline information extraction workflow:

- Captures and filters the desired information and gathers them in a easy format (json and csv files) for analysis
- An “entry”/“row” per report, gathered by “Volcano ID”, with 22 “variables”/“columns”

	colour_code	fl_eruption_L2	no_id_satelite	fl_eruption_L1	file_name	id	confirmed	info_source	issued_time	is
0502-09	yellow		-1		Darwin/2006/IDD41300.200607130420.txt	0502-09	1	RVO/Sat/Imagery	0420	2006
	yellow		-1		Darwin/2006/IDD41300.200607140356.txt	0502-09	1	RVO/MTSAT-1R/TERRA/MODIS	0356	2006
0502-08	orange		0	140	Darwin/2012/IDD41290.201207101643.txt	0502-08	1	MTSAT-2	1643	20120710
	orange		0	140	Darwin/2012/IDD41290.201207102229.txt	0502-08	1	MTSAT-2	2229	20120710
	orange		2	140	Darwin/2012/IDD41290.201207110417.txt	0502-08	1	MTSAT-2	0417	20120711
	red		0	400	Darwin/2012/IDD41295.201205030543.txt	0502-08	1	MTSAT	0542	20120503
	red		0	450	Darwin/2012/IDD41295.201205030608.txt	0502-08	1	MTSAT-2	0607	20120503
	orange		0	450	Darwin/2012/IDD41295.201205030722.txt	0502-08	1	MTSAT-2	0721	20120503
	colour_code	fl_eruption_L2	no_id_satelite	fl_eruption_L1	file_name	id	confirmed	info_source		
		0		Anchorage/FVAK21PAWU_1701121900.txt	311300	1		SATELLITE		
		-1		Anchorage/FVAK21PAWU_1701150130.txt	311300	1		HIMAWARI/GOES/SEISMICITY		
		0		Anchorage/FVAK21PAWU_1701181405.txt	311300	1		HIMAWARI/GOES/AVO/PILOT/REPORT/LIG		
		0		Anchorage/FVAK21PAWU_1701181910.txt	311300	1		HIMAWARI/GOES		
		0		Anchorage/FVAK21PAWU_1701190140.txt	311300	1		GOES/MODEL/DATA		
		2		Anchorage/FVAK21PAWU_1701190635.txt	311300	1		GOES/POES/		
		0		Anchorage/FVAK21PAWU_1701201515.txt	311300	1		HIMAWARI/GOES/POES/AVO/PIREPS		
		0		Anchorage/FVAK21PAWU_1701201950.txt	311300	1		HIMAWARI/GOES/POES		
		2		Anchorage/FVAK21PAWU_1701210140.txt	311300	1		HIMAWARI/GOES/POES		
		0		Anchorage/FVAK21PAWU_1701221805.txt	311300	1		HIMAWARI/GOES		

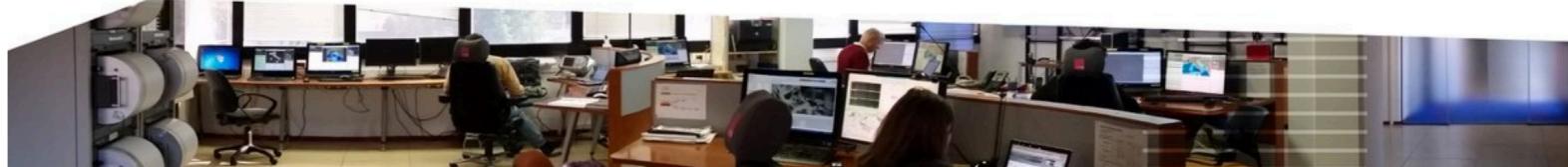
Seismology

Python script for downloading events and parsing html to select Location and Magnitude

- INGV- CENTRONEAZIONALE - TERREMOTI
<http://cnt.rm.ingv.it/en>

INGV | THE INSTITUTE | ENVIRONMENT | VOLCANOES | EARTHQUAKES | HIGHLIGHTS AND ACTIVITIES | MEDIA

 INGVCENTRONEAZIONALETERREMOTI 



EARTHQUAKE LIST INSTRUMENTS SCIENTIFIC PRODUCTS ▾ SITE GUIDE CONTACT  

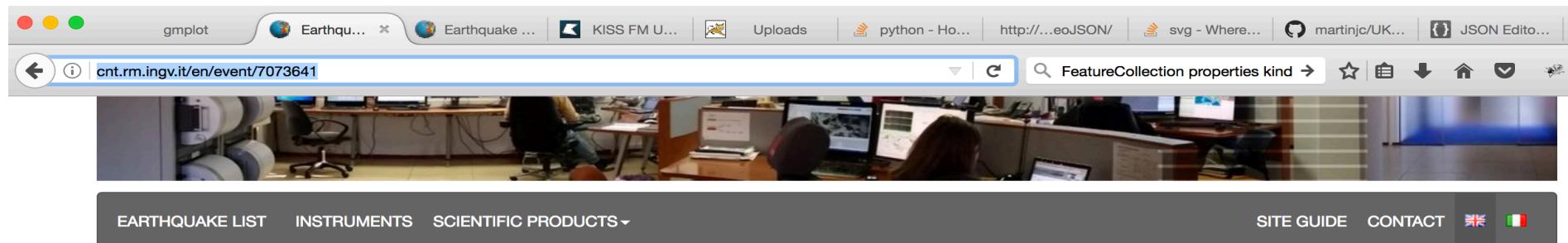
▼ Last 7 days ▼ Magnitude 2+ ▼ World Custom Search Map Export list ▾

Showing earthquakes from 1 to 50 of 220 found (Sort Time Descending)

Date and Time (UTC) ⓘ	Magnitude ↴	Region ↴	Depth ↴	Latitude	Longitude
2017-02-09 14:14:41	3.5	Perugia	7	42.68	12.70
2017-02-09 13:25:22	2.0	L'Aquila	10	42.50	13.41
2017-02-09 13:11:52	2.0	Siena	7	43.28	11.24
2017-02-09 13:00:18	2.2	Macerata	9	42.98	13.03
2017-02-09 12:37:24	2.1	Rieti	12	42.60	13.22
2017-02-09 10:50:25	2.1	L'Aquila	11	42.38	13.37

Seismology

- INGV- CENTRONAZIONALE - TERREMOTI
<http://cnt.rm.ingv.it/en>



The screenshot shows a web browser window with the URL <http://cnt.rm.ingv.it/en/event/7073641>. The page displays information about a magnitude 6.0 earthquake that occurred on August 24, 2016, at 01:36:32 UTC in the Rieti region. The interface includes a navigation bar with links for EARTHQUAKE LIST, INSTRUMENTS, SCIENTIFIC PRODUCTS, SITE GUIDE, CONTACT, and flags for English and Italian. Below the navigation bar is a large image of a control room with multiple computer monitors displaying seismic data. The main content area features tabs for Event data, Seismicity and Hazard, Impact, Locations and Magnitudes (which is currently selected), Focal mechanism, and Download. Under the Locations and Magnitudes tab, there is a table titled "Locations history" with columns for Type, Date and Time (UTC), Latitude, Longitude, Magnitude, Depth (km), Published time (UTC), Author, and Location ID. The table lists four entries corresponding to the four seismic events mentioned in the text.

Type	Date and Time (UTC)	Latitude	Longitude	Magnitude	Depth (km)	Published time (UTC)	Author	Location ID
Rev 100	2016-08-24 01:36:32	42.71	13.22	ML 6.0	4	2016-08-24 01:53:18	Sala Sismica INGV-Roma	26952071
Rev 501	2016-08-24 01:36:32	42.71	13.22	Mw 6.0	5	2016-08-24 03:14:39	TDMT-INGV Revised	26962871
Rev 1000 ★	2016-08-24 01:36:32	42.7	13.23	ML 6.0	8	2016-08-31 06:40:09	Bollettino Sismico Italiano INGV	27629391
Rev 501	2016-08-24 01:36:32	42.71	13.22	Mw 6.0	5	2016-09-23 11:19:11	TDMT-INGV Revised	29420591

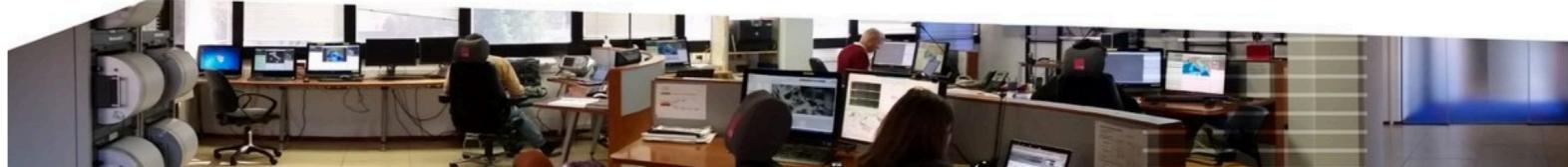


Seismology

- INGV- CENTRONAZIONALE - TERREMOTI
<http://cnt.rm.ingv.it/en>

INGV | THE INSTITUTE | ENVIRONMENT | VOLCANOES | EARTHQUAKES | HIGHLIGHTS AND ACTIVITIES | MEDIA

INGV CENTRONAZIONALE TERREMOTI



EARTHQUAKE LIST INSTRUMENTS SCIENTIFIC PRODUCTS ▾ SITE GUIDE CONTACT

▼ Last 7 days ▼ Magnitude 2+ ▼ World

Custom Search Map Export list ▾

Showing earthquakes from 1 to 50 of 220 found (Sort Time Descending)

Date and Time (UTC) ⓘ	Magnitude ↴	Region ↴	Depth ↴	Latitude	Longitude
2017-02-09 14:14:41	3.5	Perugia	7	42.68	12.70
2017-02-09 13:25:22	2.0	L'Aquila	10	42.50	13.41
2017-02-09 13:11:52	2.0	Siena	7	43.28	11.24
2017-02-09 13:00:18	2.2	Macerata	9	42.98	13.03
2017-02-09 12:37:24	2.1	Rieti	12	42.60	13.22
2017-02-09 10:50:25	2.1	L'Aquila	11	42.38	13.37

Seismology

web services

Starttime
Endtime
Magnitude
Area



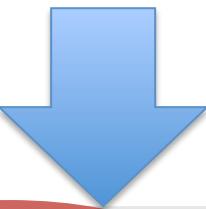
webservice.ingv.it/fdsnws/event/1/query?starttime=2017-02-02T00%3A00%3A00&endtime=2017-02-09T23%3A59

FeatureCollection properties kind

#EventID	Time	Latitude	Longitude	Depth/km	Author	Catalog	Contributor	ContributorID	Magnitude	Type	Author	EventLocationName
1324660	2017-02-02T03:16:42.720000	46.1562	12.3433	9.2	SURVEY-INGV		ML	2.3	--	Belluno		
1324821	2017-02-02T05:00:56.030000	42.9988	13.0307	5.0	SURVEY-INGV		ML	2.7	--	Macerata		
1324931	2017-02-02T05:45:35.480000	42.9967	13.0302	6.1	SURVEY-INGV		ML	2.3	--	Macerata		
13249731	2017-02-02T06:23:16.720000	38.2257	14.5758	10.0	SURVEY-INGV		ML	2.1	--	Costa Siciliana nord-orientale (Messina)		
13252031	2017-02-02T08:05:18.320000	42.7392	13.257	10.5	SURVEY-INGV		ML	2.0	--	Ascoli Piceno		
13253771	2017-02-02T09:38:16.660000	43.017	13.0368	6.8	SURVEY-INGV		ML	2.1	--	Macerata		
13254331	2017-02-02T10:03:59.450000	42.6158	13.2908	9.8	SURVEY-INGV		ML	2.0	--	Rieti		
13255851	2017-02-02T11:07:03.550000	42.7828	13.6077	25.1	SURVEY-INGV		ML	2.1	--	Ascoli Piceno		
13256911	2017-02-02T12:01:53.930000	42.5052	13.3073	14.3	SURVEY-INGV		ML	2.0	--	L'Aquila		
13256981	2017-02-02T12:03:48.000000	42.5815	13.2247	13.1	SURVEY-INGV		ML	2.0	--	L'Aquila		
13264031	2017-02-02T17:40:25.850000	42.6007	13.3123	10.2	SURVEY-INGV		ML	2.0	--	Rieti		
13265301	2017-02-02T19:03:08.100000	42.7157	13.2528	13.4	SURVEY-INGV		ML	2.2	--	Rieti		
13266611	2017-02-02T20:05:48.420000	42.6553	13.2782	12.4	SURVEY-INGV		ML	2.1	--	Rieti		
13268161	2017-02-02T21:42:49.430000	42.5783	13.2552	12.6	SURVEY-INGV		ML	2.4	--	L'Aquila		
13268511	2017-02-02T22:04:25.250000	42.7948	13.1952	11.0	SURVEY-INGV		ML	2.0	--	Perugia		
13270371	2017-02-02T23:44:12.450000	41.2193	14.8628	11.0	SURVEY-INGV		ML	2.2	--	Benevento		
13271781	2017-02-03T01:19:31.440000	42.4782	13.2857	13.8	SURVEY-INGV		ML	2.0	--	L'Aquila		
13272071	2017-02-03T01:32:08.060000	42.7917	13.316	16.8	SURVEY-INGV		ML	2.4	--	Ascoli Piceno		
13272231	2017-02-03T01:37:06.860000	46.0547	13.0658	9.7	SURVEY-INGV		ML	2.3	--	Udine		
13272491	2017-02-03T01:49:15.420000	42.5542	13.2298	10.0	SURVEY-INGV		ML	2.2	--	L'Aquila		
13273251	2017-02-03T02:26:30.030000	38.6832	11.7805	10.0	SURVEY-INGV		ML	3.2	--	Tirreno Meridionale (MARE)		
13274401	2017-02-03T03:32:59.300000	42.9932	13.0213	8.9	SURVEY-INGV		ML	3.0	--	Macerata		
13274561	2017-02-03T03:47:55.970000	42.9902	13.0218	6.4	SURVEY-INGV		Mw	4.0	--	Macerata		
13275241	2017-02-03T03:53:28.620000	42.9912	13.0352	6.6	SURVEY-INGV		ML	2.8	--	Macerata		
13275291	2017-02-03T03:54:25.670000	42.9997	13.0323	5.8	SURVEY-INGV		ML	3.0	--	Macerata		
13275591	2017-02-03T03:59:34.880000	42.9902	13.0332	5.4	SURVEY-INGV		ML	2.2	--	Macerata		
13275711	2017-02-03T04:01:16.260000	42.9858	13.0302	6.3	SURVEY-INGV		ML	2.2	--	Macerata		
13276121	2017-02-03T04:10:05.430000	42.9895	13.025	5.7	SURVEY-INGV		Mw	4.2	--	Macerata		
13276311	2017-02-03T04:12:21.700000	40.1008	15.9873	10.6	SURVEY-INGV		ML	2.4	--	Potenza		
13276371	2017-02-03T04:12:48.900000	43.0083	13.0222	4.0	SURVEY-INGV		ML	3.1	--	Macerata		
13276411	2017-02-03T04:14:00.300000	42.9903	13.0328	6.1	SURVEY-INGV		ML	2.5	--	Macerata		
13276731	2017-02-03T04:19:28.070000	42.9918	13.035	7.0	SURVEY-INGV		ML	3.3	--	Macerata		
13276821	2017-02-03T04:21:12.830000	42.9933	13.0312	6.1	SURVEY-INGV		ML	2.5	--	Macerata		
13277001	2017-02-03T04:24:08.820000	43.0088	13.0223	8.5	SURVEY-INGV		ML	2.1	--	Macerata		
13277041	2017-02-03T04:26:22.430000	42.9927	13.0293	5.6	SURVEY-INGV		ML	2.0	--	Macerata		
13279061	2017-02-03T05:11:22.560000	42.9842	13.0423	5.5	SURVEY-INGV		ML	2.8	--	Macerata		
13279931	2017-02-03T05:34:45.860000	42.9862	13.0333	7.6	SURVEY-INGV		ML	2.4	--	Macerata		
13280161	2017-02-03T05:40:34.310000	43.0025	13.0362	7.5	SURVEY-INGV		Mw	3.8	--	Macerata		
13280531	2017-02-03T05:45:15.290000	42.9903	13.0377	7.6	SURVEY-INGV		ML	2.3	--	Macerata		
13280601	2017-02-03T05:46:32.830000	42.9977	13.0222	8.4	SURVEY-INGV		ML	2.2	--	Macerata		
13280671	2017-02-03T05:48:03.100000	43.0013	13.0372	7.4	SURVEY-INGV		ML	2.1	--	Macerata		
13280811	2017-02-03T05:49:16.800000	42.986	13.047	5.1	SURVEY-INGV		ML	2.8	--	Macerata		
13281211	2017-02-03T05:54.46390000	43.0047	13.0367	7.5	SURVEY-INGV		Mt	2.1	--	Macerata		

event IDs

event IDs



Seismology

HTML

Screenshot of a web browser showing the INGV Seismology HTML page for an earthquake event.

The browser tab shows the URL: cnt.rm.ingv.it/en/event/7073641.

A red circle highlights the URL bar, and a red oval highlights the "Locations history" table below.

The page title is "Earthquake with magnitude of 6.0 on date 24-08-2016 and time 01:36:32 (UTC) in region Rieti".

The navigation menu includes: EARTHQUAKE LIST, INSTRUMENTS, SCIENTIFIC PRODUCTS, SITE GUIDE, CONTACT, and flags for English and Italian.

The "Locations and Magnitudes" tab is active.

Locations history

Type	Date and Time (UTC)	Latitude	Longitude	Magnitude	Depth (km)	Published time (UTC)	Author	Location ID
Rev 100	2016-08-24 01:36:32	42.71	13.22	ML 6.0	4	2016-08-24 01:53:18	Sala Sismica INGV-Roma	26952071
Rev 501	2016-08-24 01:36:32	42.71	13.22	Mw 6.0	5	2016-08-24 03:14:39	TDMT-INGV Revised	26962871
Rev 1000 ★	2016-08-24 01:36:32	42.7	13.23	ML 6.0	8	2016-08-31 06:40:09	Bollettino Sismico Italiano INGV	27329391
Rev 501	2016-08-24 01:36:32	42.71	13.22	Mw 6.0	5	2016-09-23 11:19:11	TDMT-INGV Revised	29420591

Code - GitHub

- Simple python code for downloading data in memory – I used some html libraries for parsing HTML – Everything is streamed – Not intermediate files
- It could be easily be transformed into a data-pipeline workflow to download/parser several events in parallel – not need for this now –
- [https://github.com/rosafilgueira/
Data_Science_Projects_BGS/tree/master/
Seismology_Example/download_cnt.py](https://github.com/rosafilgueira/Data_Science_Projects_BGS/tree/master/Seismology_Example/download_cnt.py)