# IIP InsightCloud

## Datasource Documentation

**Version 0.4 Beta**

**June 09, 2015**

## Contents

**Revision Table**

| *Version Number* | *Date* | *Description* | *Authors* |
|---|---|---|---|
| 0.3 Beta | 05/27/2015 | First Draft for initial source list | DigitalGlobe |
| 0.4 Beta | 06/09/2015 | Second Draft updated to include new and updated sources | DigitalGlobe |
| | | | |
| | | | |

# 1 Scope

## 1.1 Identification

This document applies to InsightCloud Version 0.4 Beta Dated 06/09/2015.

## 1.2 Document Overview

The purpose of this document is to:

- Identify the terrain sources and restrictions.
- Identify the social media and vector types used in the software.
- Identify the sources for each social media and vector type, and any restrictions.
- Define Anthrometer attributes.

# 2 Terrain

Terrain applications run analytics using Digital Elevation Models (DEMs) and various user inputs through algorithms to return different analytic results.

## 2.1 Terrain Source List

| Source | Source Name | Coverage |
|---|---|---|
| SRTM3 | SRTM 90M | Global |
| ASTER1 | ASTER 30M | Global |
| VRICON2M | VRICON 2M | Sporadic |

## 2.2 VRICON 2M Coverage List

Port Lincoln - Australia
Camiri - Bolivia
Rio de Janeiro - Brazil
Dorchester – England
Dorset - England
Portsmouth – England
Southampton - England

Euskirchen - Germany
Megara - Greece
Nea Peramos – Greece
Al Asad Airbase (alt. Al Anbar) - Iraq
Nairobi – Kenya
Monrovia - Liberia
Tripoli – Libya
Lagos – Nigeria
Yongbyon (alt. Nyongbyon) - North Korea
Boden - Sweden
Aleppo - Syria
Ar Raqqah - Syria
Salamiya – Syria
Gombe National Park (alt. Kigoma) - Tanzania
Mwandiga – Tanzania
Entebbe Airport - Uganda
Juneau - Alaska, USA
Al Mukalla – Yemen
Ghayl Ba Wazir – Yemen
Sompe - Southern Zambia

### 2.3 Restrictions

| Source | Terrain Analytic Name | Terrain Analytic API Call Name | Restrictions |
|---|---|---|---|
| SRTM3 | Comms RFP | Comms | Outer Radius <= 30000 |
| | Cumulative Viewshed | massvs | 7000 x 7000 |
| | HLZ Touchdown | hlz | SRTM3 not used in hlz |
| | Sightline | sightline | Outer Radius <= 30000 |
| | Sitescan | sitescan | N/A |
| ASTER1 | Comms RFP | Comms | Outer Radius <= 10000 |
| | Cumulative Viewshed | massvs | 7000 x 7000 |
| | HLZ Touchdown | hlz | ASTER1 not used in hlz |
| | Sightline | sightline | Outer Radius <= 10000 |
| | Sitescan | sitescan | N/A |
| VRICON2M | Comms RFP | Comms | Outer Radius <= 3500<br>*Note: 3500 takes 3+ minutes to return results* |

| Source | Terrain Analytic Name | Terrain Analytic API Call Name | Restrictions |
|---|---|---|---|
| | | | *Note: Small radii around 721 return results quickly* |
| | Cumulative Viewshed | massvs | VRICON2M not used in massvs |
| | HLZ Touchdown | hlz | Aoi bounding box should not exceed 20000 x 20000, or the footprint of the data, whichever is smaller<br>*Note: Large aois will take 3-4 + minutes to return* |
| | Sightline | sightline | Outer Radius <= 7500<br>*Note: 7500 takes 1+ minutes to return results*<br>*Note: Small radii around 721 return results quickly* |
| | Sitescan | sitescan | VRICON2M not used in sitescan |

# 3 Social Media

The InsightCloud Social Media widget provides the analyst with a visual output of up-to-the-second geolocated social media from Twitter and RSS feeds, based on user-defined queries.

## 3.1 Twitter

### 3.1.1 Twitter Country Coverage

| Country Name | ISO Code |
|---|---|
| Afghanistan | AF |
| Algeria | DZ |
| Armenia | AM |
| Benin | BJ |
| Botswana | BW |
| Brazil | BR |
| Burkina Faso | BF |
| Cameroon | CM |

| Country Name | ISO Code |
|---|---|
| Central African Republic | CF |
| Chad | TD |
| Colombia | CO |
| Djibouti | DJ |
| Georgia | GE |
| Ghana | GH |
| Guinea | GN |
| Honduras | HN |
| Indonesia | ID |
| Iran | IR |
| Iraq | IQ |
| Ivory Coast | CI |
| Jordon | JO |
| Kenya | KE |
| Lebanon | LB |
| Liberia | LR |
| Libya | LY |
| Malaysia | MY |
| Maldives | MV |
| Mali | ML |
| Mauritania | MR |
| Mexico | MX |
| Morocco | MA |
| Niger | NE |

| Country Name | ISO Code |
|---|---|
| Nigeria | NG |
| North Korea | KP |
| Pakistan | PK |
| Peru | PE |
| Philippines | PH |
| Sierra Leone | SL |
| Somalia | SO |
| Sudan | SD |
| Syria | SY |
| Thailand | TH |
| Togo | TG |
| Trinidad and Tobago | TT |
| Tunisia | TN |
| Turkey | TR |
| United States of America | US |
| Venezuela | VE |
| Yemen | YE |

#### 3.1.2 Twitter Available Timeframe

Twitter Ingest from: Nov 2013 – current date and time

### 3.2 RSS

#### 3.2.1 RSS Source List

| Source | Category | URL |
|---|---|---|
| All Africa | AFRICA | http://allafrica.com/tools/headlines/rdf/latest/headlines.rdf |
| BBC | AFRICA | http://feeds.bbci.co.uk/news/world/africa/rss.xml |

| Source | Category | URL |
|---|---|---|
| The Hindu | ASIA | http://www.thehindu.com/news/national/?service=rss |
| The Hindu | ASIA | http://www.thehindu.com/news/international/south-asia/?service=rss |
| Straits Times | ASIA | http://straitstimes.com.feedsportal.com/c/32792/f/640960/index.rss |
| Yonhap News | ASIA | http://yonhapnews.feedsportal.com/c/35025/f/647120/index.rss |
| Oil and Gas Journal | BUSINESS | http://feeds.feedburner.com/latest-news-ogj |
| CNN | CRIME | http://rss.cnn.com/rss/cnn_crime.rss |
| Copernicus EMS | DISASTER | http://feeds.feedburner.com/CopernicusEMSMappingRushModeActivations |
| Disaster Charter | DISASTER | http://www.disasterscharter.org/DisasterCharter/RssFeed?articleType=activation&locale=en_US&companyId=1&comunityId=10729 |
| Disaster Report | DISASTER | http://www.disaster-report.com/feeds/posts/default |
| EMERCOM of Russia | DISASTER | http://en.mchs.ru/news/rss/ |
| GDACS | DISASTER | http://www.gdacs.org/xml/rss.xml |
| Guardian | DISASTER | http://www.theguardian.com/world/natural-disasters/rss |
| PDC | DISASTER | http://d2mxabrykbl1km.cloudfront.net/feed.xml |
| RSOE-EDIS | DISASTER | http://feeds.feedburner.com/RsoeEdis-EmergencyAndDisasterInformation |
| RSOE-EDIS | DISASTER | http://feeds.feedburner.com/rsoeEdis-FaaAirplaneAccidentAndIncidentInformation |
| RSOE-EDIS | DISASTER | http://feeds.feedburner.com/RsoeEdis-TropicalStormInformations |
| RSOE-EDIS | DISASTER | http://feeds2.feedburner.com/RsoeEdis-TsunamiInformation |
| RSOE-EDIS | DISASTER | http://feeds.feedburner.com/RsoeEdis-VolcanoMonitoring |
| Topix | DISASTER | http://www.topix.com/rss/news/natural-disasters |
| FEMA | GOVERNMENT | http://www.fema.gov/news/recentnews_rss.fema |
| Naval Open Source Intelligence | GOVERNMENT | http://nosint.blogspot.com/feeds/posts/default |

| Source | Category | URL |
|---|---|---|
| BBS | MIDDLE EAST | http://feeds.bbci.co.uk/news/world/middle_east/rss.xml |
| Space Daily | SPACE | http://www.spacedaily.com/spacedaily.xml |
| Space Flight Now | SPACE | http://www.spaceflightnow.com/sfnnews.xml |
| Space War | SPACE | http://www.spacewar.com/spacewar.xml |
| Terra Daily | SPACE | http://www.terradaily.com/terradaily.xml |
| Al Jazeera | TOP STORIES | http://www.aljazeera.com/Services/Rss/?PostingId=2007731105943979989 |
| AP | TOP STORIES | http://hosted.ap.org/lineups/TOPHEADS.rss?SITE=AP&SECTION=HOME |
| BBC | TOP STORIES | http://feeds.bbci.co.uk/news/rss.xml |
| CNN | TOP STORIES | http://rss.cnn.com/rss/cnn_topstories.rss |
| CNTV | TOP STORIES | http://english.cntv.cn/service/rss/0/index.xml |
| Google | TOP STORIES | https://news.google.com/news/feeds?pz=1&cf=all&ned=us&hl=en&output=rss |
| REUTERS | TOP STORIES | http://feeds.reuters.com/reuters/topNews |
| Sydney Morning Herald | TOP STORIES | http://feeds.smh.com.au/rssheadlines/national.xml |
| Yonhap News | TOP STORIES | http://yonhapnews.feedsportal.com/c/35025/f/647119/index.rss |
| CNN | US | http://rss.cnn.com/rss/cnn_us.rss |
| AP | WORLD | http://hosted.ap.org/lineups/WORLDHEADS.rss?SITE=AP&SECTION=HOME |
| CNN | WORLD | http://rss.cnn.com/rss/cnn_world.rss |
| BBC | WORLD | http://feeds.bbci.co.uk/news/world/rss.xml |

#### 3.2.2 RSS Available Timeframe

RSS Ingest from: Nov 2013 – current date and time
*Note: Not all datasources were added at the same time.*

### 3.3 Restrictions

Kernel Density restricts pageSize to 100000.

# 4 UVI

The Unified Vector Index (UVI) provides the analyst with a tool for visually mapping all available vectors within a given aoi. Analysts may then refine the resulting vectors into a desired selection. These quickly polled and returned arrays of points, polylines, and polygons may then be used in further analytical studies of the area.

## 4.1 Restrictions

Querying the index without the use of paging is restricted to 1000 vector items.
Querying the index through paging is unrestricted for the number of items returned.

## 4.2 Source List

| Source Name | Coverage |
|---|---|
| ACLED | Africa |
| Anthrometer | Global |
| Gazetteer | Global |
| GDB | US |
| HGIS | Nigeria; Togo; Mali; Benin; Berkina Faso; Sierra Leone; Ghana; Guinea; Liberia; Niger; Ivory Coast |
| OSM | Global |
| SETD | Syria |
| Tomnod | Nepal; Lucknow, India; Tibet (north of Nepal, south of Bogkamba) |

## 4.3 Anthrometer Attribute Definitions

| Level | Attr. Name | Description |
|---|---|---|
| Top Level | Geometry | The GeoJSON polygon representing this bin's geohash |
| properties | itemType | The date at the start of this bin's time period |
| | itemDate | The start of the time period for this geotemporal bin (all bins are currently 15 minutes long) |
| properties.attributes (general attributes) | TimeBucket | Same as properties.itemDate |
| properties.attributes (volume analysis | VolumeActualValue | The actual volume of tweets in this geotemporal bin |

| Level | Attr. Name | Description |
|---|---|---|
| attributes) | VolumeExpectedMin / VolumeExpectedMax | The range of expected values based on the model; the model is currently trained on the three immediately prior weeks of historic data |
| | GeohashType | The name of the model variation that was used to analyze this bin (currently always "low-volume") |
| | AlertLevel | A value indicating how significant the deviation of the actual tweet volume is from the expected range; a value twice as high is approximately twice as unusual; a value of at least 1.0 is a "significant" deviation |
| | AlertType | A simple categorization of the values in AlertLevel; an AlertLevel of 0.0 yields an AlertType of "none," an AlertLevel less than 1.0 yields an AlertType of "watch," and an AlertLevel at least 1.0 yields an AlertType of "warning" |
| properties.attributes (sentiment analysis attributes) | MedianSentimentMedian | An "average" level of sentiment in this bin (to be precise, it is the median of the median of the non-0.5 positive sentiment values from each training time bucket in this geohash) |
| | SentimentMAD | An "average" difference between MedianSentimentMedian and the individual sentiment medians from the training time buckets |
| | SentimentMedian | The median of the non-0.5 positive sentiment values from the tweets in this bin |
| | SentimentDeviation | The level of deviation of this bin's SentimentMedian from the MedianSentimentMedian; this value uses SentimentMAD for limited "normalization," but the resultant value should not be used as a ratio value (i.e. it's only useful to say that a higher value is more unusual); a value of at least 10.0 is a "significant" deviation |

*Note: There will only be entries for geotemporal bins if the AlertLevel is at least 1.0 or the SentimentDeviation is at least 10.0.*