CrowdSignals Platform

Description of Sensor Data Types Collected

User's Reference Document

02/10/2016

P.O. Box 4452 91 S Jackson S, Seattle, WA 98104

EMAIL
info@AlgoSnap.com

WEB
www.AlgoSnap.com

ALGOSNAP

CrowdSignals Platform Data Types Collected

Purpose:

The purpose of this document is to list the different types of data collected by the CrowdSignals Platform and anonymization techniques applied (whenever necessary or selected by customers) to the data collected from participants all across the United States.

Overview

In the following sections we describe each collected data type in detail as well as some potential research use cases and steps that are taken to anonymize the data. All data types include the following fields as metadata:

User ID	The unique ID of the subject (e.g., "138")
Type	The data type (e.g., "Bluetooth", "accelerometer", "SMS")
Device Type	The type of device (e.g., smartphone, smartwatch, tablet)
Device ID	A unique ID for each user's device. This is be mapped to the IMEI of the device but not made available.
Start	The timestamp in nanoseconds associated with the start of the window of data samples
End	The timestamp in nanoseconds associated with the end of the window of data samples
Timezone	The timezone in which the data was logged.
SW Name	The name of the software collecting the data
SW Version	The version number of the software collecting the data
Timestamps	A list of timestamps in nanoseconds indicating the time at which each data record sample was created.

Anonymization: The device ID (e.g., device IMEI) will be anonymized via one-way hash so that specific devices cannot be identified.

Example Record Data (Bluetooth in this case):

```
{
   'user_id':0,
   'type':'bluetooth',
   'device_type':'smartphone',
   'device_id':'358239056736392'
   'start':1454465357000000000,
   'end':1454465387000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'timestamps':[
      1454465363201000000,
      1454465367101000000,
      …
      1454465384244000000
   ]
}
```

Geo-Location and Radios

We're collecting a variety of location and radio data which is useful in studies of mobility and mobility patterns as well as mobile network usage and performance.

Bluetooth Record

A scan of Bluetooth-enabled devices within range of the subject's smartphone or smartwatch including the following fields:

Address	A list containing the hardware address of the scanned devices
Name	A list containing the human-readable name of the scanned devices (e.g., "Bob's iPhone"). This field can be modified by a device owner and can contain information such as their name.
RSSI	A list containing the received signal strength of each scanned device
Timestamps	A list containing the timestamps in nanoseconds at which each scanned device was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional Anonymization: We apply a two-step anonymization process to Bluetooth records. First, on the device before Bluetooth data is uploaded, the Address and Name fields are one-way hashed to remove any uniquely identifiable and human readable information. At the end of the data collection period, all Bluetooth data will then be further anonymized by replacing all hashed results with numeric identifiers so that each Name and Address are associated with exactly one numeric identifier (e.g., "Bob's iPhone" = 18476). Also, the device id is mapped to a unique id that has nothing to do with the user device ID.

Example Record Data (not Anonymized):

```
{
   'user_id':0,
   'type':'bluetooth',
   'device_type':'smartphone',
   'device_id':'358239056736392'
   'start':1454465357000000000,
   'end':1454465387000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'address':[
      '1C:E6:2B:E6:1C:A7',
      '00:1D:86:5F:A6:6C',
       …
      '1C:1A:C0:68:DD:F9'
   ],
   'name':[
      'null','KDC-BT3**U', …, 'null',
   ],
   'timestamps':[
      1454465363201000000,1454465367101000000, …, 1454465381830000000
   ],
   'rssi':[ -94,-92, …, -90 ]
}
```

WLAN Record

A scan of WLAN access points or WLAN-enabled devices within range of the subject's smartphone or smartwatch including the following fields:

BSSID	A list containing the hardware address of the scanned access points
SSID	A list containing the human-readable name of the scanned access points (e.g., "Bob's Apartment"). These names can be edited by the WLAN access point owner and might contain place related information or ownership information such as "Starbucks-cafe" or "Apartment 411, 4123 El Camino"
Frequency	A list containing the frequency of each scanned access point
Level	A list containing the received signal strength of each scanned access point
Timestamps	A list containing the timestamps in nanoseconds at which each scanned access point was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: We apply a two-step anonymization process to WLAN records. First, on the device before WLAN data is uploaded, the BSSID and SSID fields are one-way hashed to remove any uniquely identifiable and human readable information. At the end of the data collection period, all WLAN data will then be further anonymized by replacing all hashed results with numeric identifiers so that each BSSID and each SSID are associated with exactly one numeric identifier (e.g., "Bob's Apartment" = 286732). Also, the device id is mapped to a unique id that has nothing to do with the user device ID.

Example Record Data (not Anonymized):


```
{
   'user_id':0,
   'type':'wlan',
   'device_type':'smartphone',
   'device_id':'358239056736392',
   'start':1454466087000000000,
   'end':1454466097000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'ssid':[
      'NNDS_Wifi5G',
      'noris',
      ...
      'NETGEAR'
   ],
   'bssid':[
      '00:02:6f:e2:55:94',
      '40:16:7e:31:4f:fc',
      ...
      '28:10:7b:2f:73:ae'
   ],
   'level':[
      -87.0,
      -88.0,
```

```
        ...
        -77.0
    ],
    'frequency':[
       5180,
       5785,
       ...
       2432
    ],
    'timestamps':[
       1454466088776381861,
       1454466089578133194,
       ...
       1454466092851481090
    ]
}
```

GSM Record

A scan of GSM cell towers within range of the subject's smartphone or smartwatch including the following fields:

MCC	A list containing the mobile country code of scanned towers
MNC	A list containing the mobile network code of scanned towers
LAC	A list containing the local area code, tracking area code, or network ID of scanned towers
CID	A list containing the cell IDs of each scanned tower
Timestamps	A list containing the timestamps in nanoseconds at which each scanned tower was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: We apply a two-step anonymization process to GSM records. First, on the device before GSM data is uploaded, the MCC, MNC, LAC, and CID fields are one-way hashed to remove any information that could be used to map the reading back to a specific tower. At the end of the data collection period, all GSM data will then be further anonymized by replacing all hashed results with numeric identifiers so that each MCC, MNC, LAC, and CID are associated with exactly one numeric identifier. Also, the device id is mapped to a unique id that has nothing to do with the user device ID.

Example Data Record:

```
{
    'user_id':0,
    'type':gsm,
    'device_type':'smartphone',
    'device_id':'358239056736392',
    'start':1454466087000000000,
    'end':1454466097000000000,
    'timezone':-8,
    'sw_name':'crowdsignals.algosnap.com',
    'sw_version':'1.0(1)',
```

```
   'mcc': [
      "310"
   ],
   'mnc': [
      "410"
   ],
   'lac': [
      37145
   ],
   'cid': [
      108011274
   ],
   'timestamps': [
      1447619526545
   ]
}
```

Location Record

A window of location readings provided by Android's location service.

Latitude	A list containing the latitude for each reading
Longitude	A list containing the longitude for each reading
Accuracy	A list containing the estimated accuracy for each reading
Speed	A list containing the speed at which the device was moving
Bearing	A list containing the bearing of the device while moving
Provider	A list containing the name of the specific service or sensor providing the location information
Timestamps	A list containing the timestamps in nanoseconds at which each reading was taken

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: We apply a subject-assisted process to degrade and anonymize location records. First we apply state-of-the-art spatial data mining algorithms to extract "significant places" from the location data. These are geographic areas where the subject has spent considerable time, they are specified by a centroid and a radius (i.e., a circle on a map). Then at the end of the data collection period we ask the subject to review each extracted place on a web interface (i.e., a Google Map interface with an overlay showing each place and when it was visited). The subject will then provide a label for each place (e.g., "home", "work", "grocery") and all other location information will be removed, leaving only the following fields in the final CrowdSignals.io dataset. Also, the device id is mapped to a unique id that has nothing to do with the user device ID.

Example location anonymized record:

Place Name	A list containing the subject-provided name of each place visited
Place Category	A list containing the subject-chosen place category (e.g., restaurant, store)
Timestamps	A list containing the timestamps in nanoseconds at which each place was entered

Example Data Record:


```
{
  'user_id': 0,
  'type': 'location',
  'device_type': 'smartphone',
  'device_id': '352622063881655',
  'start': 0,
  'end': 0,
  'timezone': -8,
  'sw_name': 'crowdsignals.algosnap.com',
  'sw_version': '1.0(1)',
  'latitude': [
    47.53992
  ],
  'longitude': [
    -122.26874
  ],
  'accuracy': [
    40.5
  ],
  'speed': [
    0
  ],
  'bearing': [
    0
  ],
  'provider': [
    'network'
  ],
  'timestamps': [
    1454522621738
  ]
}
```

Sensors

We're collecting a variety of sensor data which is useful in studies on the recognition and analysis of human activities, situations, and environments.

Accelerometer Record

A window of readings from an accelerometer

X	A list containing the acceleration on the x-axis for each reading
Y	A list containing the acceleration on the y-axis for each reading
Z	A list containing the acceleration on the z-axis for each reading
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Accelerometer readings pose a minimal privacy threat and will not be anonymized.

Example Record Data:


```
{
  'user_id':0,
  'type':'accelerometer',
  'device_type':'smartphone'
  'device_id':'358239056736392'
  'start':1454466225000000000,
  'end':1454466226000000000,
  'timezone':-8,
  'sw_name':'crowdsignals.algosnap.com',
  'sw_version':'1.0(1)',
  'x':[
    0.8988494873046875,
    0.868988037109375,
    ...
    1.002197265625
  ],
  'y':[
    5.80572509765625,
    5.8250274658203125,
    ...
    5.5736236572265625
  ],
  'z':[
    8.5299072265625,
    9.085800170898438,
    ...
    6.95703125
  ],
  'timestamps':[
```

```
    1454466225006687010,
    1454466225011863452,
    ...,
   1454466225999079585
  ]
}
```

Gyroscope Record

A window of readings from a gyroscope

X	A list containing the angular speed around the x-axis for each reading
Y	A list containing the angular speed around the y-axis for each reading
Z	A list containing the angular speed around the z-axis for each reading
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional Anonymization: Gyroscope readings pose a minimal threat and will not be anonymized.

Example Record Data:

```
{
    'user_id':0,
    'type':'gyroscope',
    'device_type':'smartwatch',
    'device_id':'IMEIUnknown',
    'start':1454458232000000000,
    'end':1454458233000000000,
    'timezone':-8,
    'sw_name':'crowdsignals.algosnap.com',
    'sw_version':'1.0(1)',
   'y':[
      0.007990118116140366,
      0.03462384268641472,
      ...
      0.06392094492912292
   ],
   'x':[
      0.0026633725501596928,
      0.010653490200638771,
      ...
      0.01598023623228073
   ],
   'z':[
      0.0026633725501596928,
      0.007990118116140366,
      ...
```

```
      0.007990118116140366
   ],
    'timestamps':[
      1454458232437025635,
      1454458232440110108,
      ...
      1454458232444964600
   ]
}
```

Magnetometer Record

A window of readings from a magnetometer

X	A list containing the magnetometer reading on the x-axis for each reading
Y	A list containing the magnetometer reading on the y-axis for each reading
Z	A list containing the magnetometer reading on the z-axis for each reading
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Magnetometer readings pose a minimal threat and will not be anonymized.

Record Example Data:

```
{
    'user_id':0,
    'type':'magnetometer'
    'device_type':'smartphone' or ‘smartwatch’,
    'device_id':'358239056736392',
    'start':1454458137000000000,
    'end':1454458138000000000,
    'timezone':-8,
    'sw_name':'crowdsignals.algosnap.com',
    'sw_version':'1.0(1)',
   'x':[
      4.82177734375,
      5.05828857421875,
      ...
      5.11474609375
   ],
   'y':[
      -26.30157470703125,
      -26.77459716796875,
      ...
      -26.83258056640625
   ],
```

```
   'z':[
      24.68414306640625,
      24.2218017578125,
      ...
      24.05548095703125
   ],
   'timestamps':[
      1454458137866906935,
      1454458137886988588,
      ...
      1454458137907468419
   ]
}
```

Heart Rate Record

A window of readings from user's heart rate data in beats per minute.

Rate	The user's heart rate in beats per minute
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Accelerometer readings pose a minimal privacy threat and will not be anonymized.

Example Record Data:

```
{
   'user_id': 0,
   'type': 'heartrate',
   'device_type': 'smartwatch',
   'device_id': 'IMEIUnknown',
   'start': 1454459245000000000,
   'end': 1454459255000000000,
   'timezone': -8,
   'sw_name': 'crowdsignals.algosnap.com',
   'sw_version': '1.0(1)',
   'rate': [ 66, 65, ..., 67 ],
   'timestamps': [
      1454459245290253334,
      1454459245417344844,
      ...
      1454459245537199114
   ]
}
```

Temperature

A window of readings for ambient temperature

Temperature	A list containing the ambient temperature readings in Celsius degrees
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Ambient temperature readings pose a minimal threat and will not be anonymized.

Example Record Data:

```
{
   'user_id':0,
   'type':temperature,
   'device_type':'smartphone',
   'start':1454466215000000000,
   'end':1454466216000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'device_id':'358239056736392'
   temperature:[
      70.1,
      70.5,
      ...
      70.5
   ],
    'timestamps':[
      1454466215032061669,
      1454466215065295432,
      ...
      1454466215099344821
   ]
}
```

Light

A window of light readings

Lux	A list containing the ambient light readings
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Light readings pose a minimal threat and will not be anonymized.

Example Record Data:


```
{
   'user_id':0,
   'type':'light',
   'device_type':'smartphone',
   'device_id':'358239056736392',
   'start':1454465796000000000,
   'end':1454465797000000000,
    'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'lux':[
      0.0
   ],
   'timestamps':[
      1454465796993392868
   ]
}
```


Pressure

A window of pressure readings

Pressure	A list containing the pressure readings in mercury millibars
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Pressure readings pose a minimal threat and will not be anonymized.

Example Record Data:


```
{
   'user_id':0,
   'type':'pressure',
   'device_type':'smartphone',
   'start':1454466215000000000,
   'end':1454466216000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'device_id':'358239056736392'
   'pressure':[
      1025.5662841796875,
      1025.5662841796875,
      ...
```

```
      1025.5662841796875
   ],
    'timestamps':[
      1454466215032061669,
      1454466215065295432,
      ...
      1454466215099344821
   ]
}
```

Humidity

A window of humidity readings

Pressure	A list containing the ambient relative humidity readings in relative humidity (%)
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Humidity readings pose a minimal threat and will not be anonymized.

Example Record Data:

```
{
   'user_id':0,
   'type':humidity,
   'device_type':'smartphone',
   'start':1454466215000000000,
   'end':1454466216000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'device_id':'358239056736392'
   humidity:[
      60.4 ,
      60.5,
      ...
      60.4
   ],
    'timestamps':[
      1454466215032061669,
      1454466215065295432,
      ...
      1454466215099344821
   ]
}
```

Proximity

A window of proximity readings

Distance	A list containing the measure of proximity as distance from the device for each reading
MaxRange	The max possible range for distance
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Proximity readings pose a minimal threat and will not be anonymized.

Example Record Data:

```
{
   'user_id':0,
   'type':'proximity'
   'device_type':'smartphone',
   'device_id':'358239056736392',
   'start':1454466075000000000,
   'end':1454466076000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'maxRange':[
      5
   ],
   'distance':[
      0
   ],
   'timestamps':[
      1454466075450543165
   ],
}
```

Microphone (Audio) Record

A window of readings from a microphone

MediaFile	The path to the MP4 file containing the raw audio.
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Raw audio data collection: Raw audio data collection will be only performed for customers who want to record raw audio data that does not include speech signals other than that of the participant and not in public places (or very controlled conditions where all people generating speech signals agree to the recording of the data).

Optional Anonymization: Audio readings are a potentially sensitive data type and will be anonymized on the device before the data is uploaded. To anonymize audio data, the device will extract and upload only a set of features from the data. The set of features is selected so that it is very difficult to reconstruct the original audio stream and any human-intelligible content it contains. Examples of these features include FFT transformation, MFCC coefficients, and min, max, avg, energy, RMS values of the signal. These signals are very similar to the features extracted by personal assistants nowadays such as SIRI, CORTANA, etc.

Example Record Data:

```
{
   'user_id':0,
   'type':'microphone',
   'device_type':'smartphone',
   'device_id':'358239056736392',
   'start':1454466047000000000,
   'end':1454466052000000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'mediaFile':'/storage/emulated/0/Guardian/MediaFiles/Audio/audio_1454466047000000000_1454466052000000000_-
8.mp4',
   'timestamps':[
      1454466047000000000
   ]
}
```

System and Network

We're collecting a variety of system and network data which is useful in studies on device and network operation and efficiency.

Battery Record

A window of readings on battery status

Level	A list containing the battery level in % for each reading
Scale	The scale on which level is recorded
Temperature	A list containing the internal battery temperature for each reading
Voltage	A list containing the voltage for each reading
Plug State	A list containing the status of the plug (i.e., plugged-in, unplugged) for each reading
Status	A list containing a string description of the battery status for each reading
Health	A list containing a string description of the battery health for each reading
Timestamps	A list containing the timestamps in nanoseconds at which each reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Anonymization: Battery readings pose a minimal threat and will not be anonymized.

Example Record Data:


```
{
    'user_id':0,
    'type':'battery',
    'device_type':'smartphone',
    'device_id':'358239056736392',
    'start':1454466137000000000,
    'end':1454466167000000000,
    'timezone':-8,
    'sw_name':'crowdsignals.algosnap.com',
    'sw_version':'1.0(1)',
   'plugged':[
      0
   ],
   'status':[
      'DISCHARGING'
   ],
   'timestamps':[
      1454466137011000000
   ],
   'health':[
      'GOOD'
   ],
   'temperature':[
      236
   ],
   'voltage':[
      3971
   ],
   'level':[
      70
   ],
   'scale':[
      100
   ]
}
```


Connectivity Record

A window of network connection events

Connected	A list containing the connection state for each event
Connecting	A list containing the an indicator as to whether the network is currently connecting for each event
Available	A list containing the availability of the network for each event
Network Type	A list containing the type of network for each event

Roaming	A list containing the roaming status for each event
SSID	A list containing the symbolic ID for the network in each event
Timestamps	A list containing the timestamps in nanoseconds at which each event was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Anonymization: The SSID field in network connection records is moderately sensitive and will be anonymized using the same two-step process (e.g., one-way hash followed by an integer mapping) with which WLAN and GSM data are anonymized.

Example Record Data:


```
{
   'user_id': 0,
   'type': 'connectivity',
   'device_type': 'smartphone',
   'device_id': '352622063881655',
   'start': 1454522618305000000,
   'end': 1454522618305000000,
   'timezone': -8,
   'sw_name': 'crowdsignals.algosnap.com',
   'sw_version': '1.0(1)',
   'connected': [ true ],
   'connecting': [ true ],
   'available': [ true ],
   'network_type': [ 'WIFI_CONNECTION' ],
   'roaming': [
      'NOT_ROAMING'
   ],
   'ssid': [
      '\'HOME-3BC2\''
   ],
   'timestamps': [
      1454522618305000000
   ]
}
```


Connection Strength

A window of GSM radio connection strength readings.

Strength	A list containing the signal strengths of each reading
Timestamps	A list containing the timestamps in nanoseconds at which each signal strength reading was observed

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Connection strength logs pose a minimal threat and will not be anonymized

Example Record Data:

```
{
   'user_id':0,
   'type':'connectionstrength',
   'device_type':'smartphone',
   'device_id':'358239056736392''end':1454458137574000000,
   'start':1454458137574000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'strength':[
      7
   ],
   'timestamps':[
      1454458137574000000
   ]
}
```

User Interaction

We're collecting a variety of user interaction logs which are useful in studies on how users actually interact with their device in a variety of contexts.

App Installation Record

A list of app install and uninstall events.

Action	A list containing the type of event
Package	A list containing the package name for the app in question (e.g., "Twitter", "Tinder", "Yelp")
Timestamps	A list containing the timestamps in nanoseconds at which each event occurred

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: App install and uninstall records pose a minimal risk and will not be anonymized

Example Record Data:

```
{
  No example because this is data type is work in progress…
}
```

App Usage Record

A snapshot of app usage at a given time.

Apps	A list of lists that contain the name of every app in usage for a given snapshot
Timestamps	A list containing the timestamps in nanoseconds at which each snapshot occurred

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Anonymization: App usage records pose a minimal risk and will not be anonymized.

Example Record Data:

```
{
  No example because this is data type is work in progress…
}
```

Camera Usage Record

A list of camera usage events.

Image File	A list of image file names resulting from camera usage
Timestamps	A list containing the timestamps in nanoseconds at which each camera event occurred

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: Camera usage records pose a minimal risk and will not be anonymized.

Example Record Data:

```
{
  No example because this is data type is work in progress…
}
```

Phone State Record

A list of phone state readings.

EventType	A list of codes indicating the type for each event
Data	A list of strings containing the phone state data for each event
Number	The phone number associated with the event, if appropriate
Timestamps	A list containing the timestamps in nanoseconds at which each event occurred

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: The Number field in phone state records is potentially sensitive and will be anonymized using the same two-step process (i.e., one-way hash followed by an integer mapping) as WLAN records.

Example Record Data:

```
{
   'user_id':0,
   'type':'phonestate',
   'device_type':'smartphone',
   'device_id':'358239056736392',
   'start':1454466198950000000,
   'end':1454466198950000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'data':[
      'DATA_ACTIVITY_NONE'
   ],
   'number':[
      'null'
   ],
   'event_type':[
      'DataActivity'
   ],
  'timestamps':[
      1454466198950000000
   ]
}
```


Screen State Record

A list of screen state readings.

State	A list of strings that indicate the screen state (e.g., on, off) for each event
Timestamps	A list containing the timestamps in nanoseconds at which each reading occurred

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: The screen state poses a minimal threat and will not be anonymized.

Example Record Data:


```
{
   'user_id':0,
   'type':'screen',
   'device_type':'smartphone',
   'device_id':'358239056736392',
   'start':1454466076509000000,
   'end':1454466076509000000,
   'timezone':-8,
   'sw_name':'crowdsignals.algosnap.com',
```

```
   'sw_version':'1.0(1)',
   'status':[
      'SCREEN_ON'
   ],
   'timestamps':[
      1454466076509000000
   ]
}
```

Social and Communication

We're collecting a variety of social and communication logs, which are useful in social science studies on patterns of communication among smartphone users.

Call Content Record

A list of call events.

Number	A list containing the phone number used in the event
CallType	A list containing the type of event (e.g., outgoing, incoming, missed)
Duration	A list containing the duration of each call in seconds
Timestamp	The timestamps in nanoseconds at which the event occurred

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Optional anonymization: The Number field in call records is potentially sensitive and will be anonymized using the same two-step process (i.e., one-way hash followed by an integer mapping) as WLAN records.

Example Record Data:

```
{
   'user_id': 0,
   'type': 'callcontent',
   'device_id': '352622063881655',
   'device_type': 'smartphone',
   'start': 1454617154387,
   'end': 1454617154387,
   'timezone': -8,
   'sw_name': 'crowdsignals.algosnap.com',
   'sw_version': '1.0(1)',
   'content_length': [
      6
   ],
   'call_type': [
      'incoming'
   ],
   'number': [
      'F7DAF72544719367FECE63C773A8EACC8EF4878626781153FB4D6A3C8865E3E4'
```

```
   ],
   call: [
      16
   ],
   'timestamp': [
      1454617154387
   ]
}
```

SMS Record

A list of SMS events.

ContentLength	A list containing the length of each SMS message in characters
NumWords	A list containing the number of words in each SMS message in characters
Address	A list containing the phone number used in the SMS event
MsgType	A list containing strings that describe the type of SMS event
EventType	A list containing the type of each SMS message (e.g., received, sent)
Timestamps	A list containing the timestamps in nanoseconds at which each event occurred

Note: for customers using our platform to collect data, anonymization is optional because there is a data license agreement where they agree to protect the confidentiality of the data, not reverse engineer it, and not transfer it to third parties.

Anonymization: The Address (phone number) field in call records is potentially sensitive and will be anonymized using the same two-step process (i.e., one-way hash followed by an integer mapping) as WLAN records.

Example Record Data:

```
{
   'user_id': 0,
   'type': 'sms',
   'device_id': '352622063881655',
   'device_type': 'smartphone',
   'start': 1454617154387,
   'end': 1454617154387,
   'timezone': -8,
   'sw_name': 'crowdsignals.algosnap.com',
   'sw_version': '1.0(1)',
   'content_length': [
      6
   ],
   'num_words': [
      2
   ],
   'address': [
      'F7DAF72544719367FECE63C773A8EACC8EF4878626781153FB4D6A3C8865E3E4'
   ],
   'msg_type': [
      '2'
   ],
   'event_type': [
      'SMSContentChanged'
```

```
   ],
   'timestamps': [
      1454617154387
   ]
}
```

User Labels

We're collecting a variety of survey data from users and user labels associated with events such as "user running" from 8am to 9am. This section explain some of those records.

Interval Label Record

A record containing a list of start and end labels provided by a user. These labels are used to provide ground truth on a particular activity, event, or situation that the user is engaged in.

Anonymization: Interval labels are explicitly created and voluntarily provided by the user and will not be anonymized.

Example Record Data:

```
{
   'user_id':0,,
   'type':'interval_label'
   'device_type':'smartphone',
   'device_id':'358239056736392',
   'sw_name':'crowdsignals.algosnap.com',
   'sw_version':'1.0(1)',
   'start':1454463944878000000,
   'end':1454464936616000000,
   'timezone':-8,
   'label':[
      'Driving'
   ],
   'label_start':[
      1454463944878000000
   ],
   'label_end':[
      1454464936616000000
   ]
}
```