Signals User Guide Documentation

Release 2.4

Stratifyd, Inc.

February 01, 2017

Signals is a powerful, yet easy-to-use augmented intelligence platform.

- Signals' *Analytics Engine* leverages unsupervised deep machine learning to identify statistically significant topics within any unstructured data.

- Signals is completely web-based. The only software required is a modern web browser such as Google Chrome, Mozilla Firefox, Microsoft Edge (Microsoft IE is not recommended).

- Signals' web interface is point-and-click. No programming knowledge is required.

This guide covers every feature and function within Signals, if you're looking for information about a specific topic, jump to a chapter in the left nav bar.

For more information about Signals or Stratifyd, please visit our website.

The chapters below go into detail about everything from getting data, to analyzing it, to sharing your newly gleaned insights.

Quick Start Guide

By the end of this section you'll have access to data of your interest, and a dashboard that allows you to analyze and navigate through structured and unstructured data.

1.1 Create an account

Begin by Creating an Account. Once you've logged in, check out some of the pre-built public dashboards in the folders on the homepage.
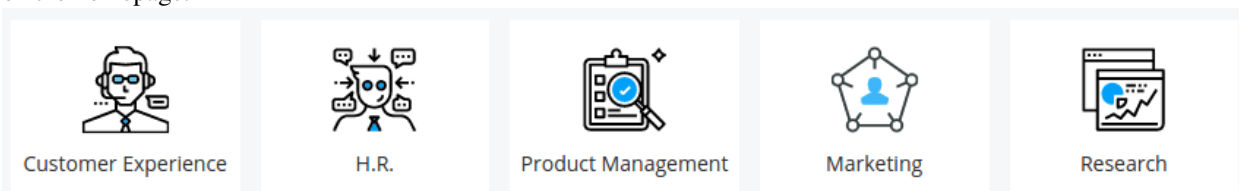
1.2 Create your first dashboard

To create your own dashboard, click "Create a new Dashboard".
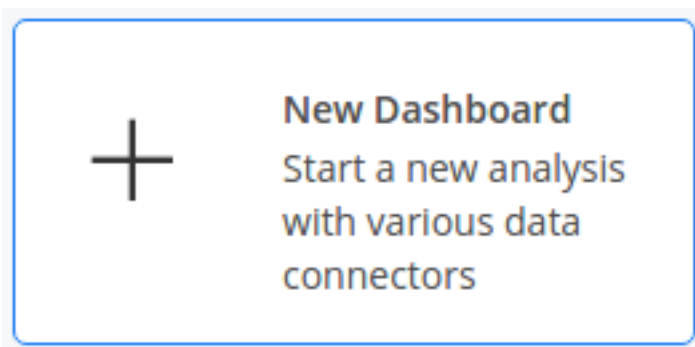
1.2.1 Connect to data

Connect to your database, upload a CSV or Excel file, or use one of our *Data Connectors* for easy access to a wealth of publicly available data. Add as many data source to your dashboard as needed.
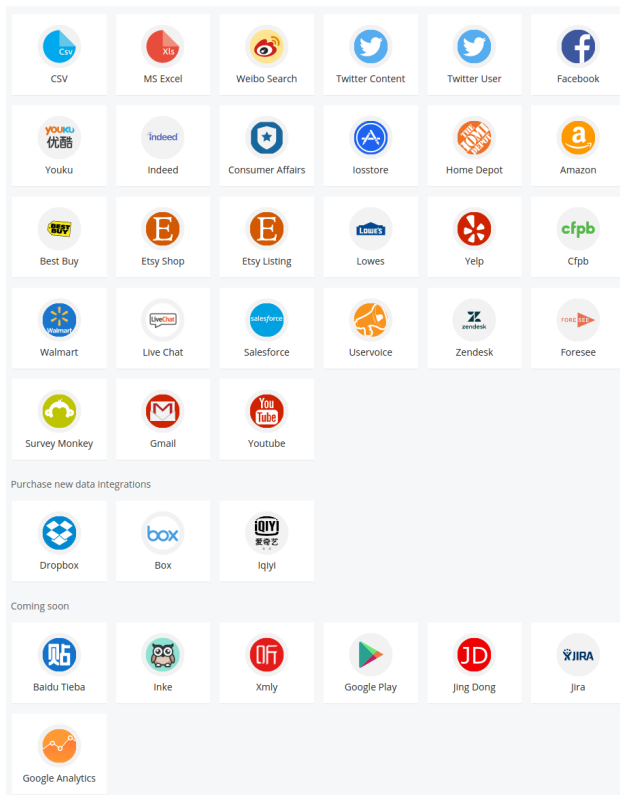
If you're bringing your own data, specify which fields contain Textual, Temporal, or Geographical information for analysis and normalization. All other fields will be brought in as pivot points and additional visualizations.
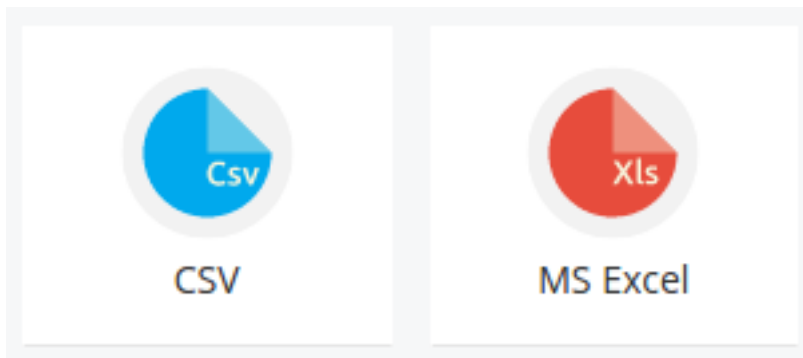
Note: See the Mapping Data guide for more details when uploading CSV or Excel Files

1.2.2 Create visualizations

Your dashboard will contain visualizations based on the type of data in your analysis by default. See the guide on *Processing Data* for information on how to interpret these visualizations. Every visualization is clickable, and clicking will apply a filter on your data based on your selection.

Now you're ready to Create Your Own Widgets using your other datapoints to home in on interesting segments within your data and discover topics and themes within your textual data.

1.2.3 Find meaningful insights

Let the visualizations guide your analysis. Click on areas of interest to drill-down into subsets of data

Click on the "Data" tab to view the original documents. If you've filtered down to a subset of your data in your dashboard, you'll see only those documents.

For more information see the chapter on *Analyzing Data*.

1.3 Share your results

Make your dashboard public or *share* with colleagues.

Connecting to Data

There are three categories of Data Connectors in Signals:

Enterprise Data Connectors are used to connect to internal databases or other 1st party data platforms in your organization.

3rd Party Data Connectors provide access to data owned by your organization but stored in 3rd party systems. Examples include salesforce, Google Analytics, and Zendesk among many others.

Public Data Connectors allow you to access data that anyone can view on the internet, but in a uniform format. We collect data from Amazon, Facebook, the Google Play store, consumeraffairs.com, and many other sites.

You can also upload a csv or excel file from your local computer.
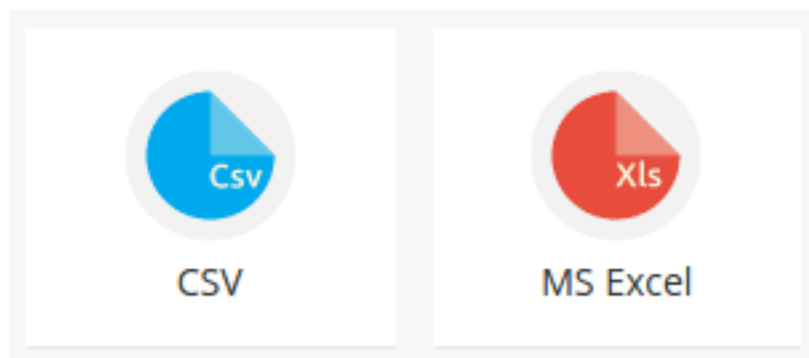
Note: Enterprise Data Connectors and some 3rd Party Data Sources require you to map fields in your data prior to uploading. See Mapping Data to learn about this process.

2.1 Enterprise Data Connectors

The Enterprise Data Connector application allows users to connect with a variety of internal data sources.

Examples include:

- Hive

- mySQL

- Microsoft SQL Server

- PostgreSQL

• Oracle SQL

2.2 3rd Party Data Connectors

3rd Party Data Connectors usually have a pop-up window that allows you to enter your credentials with the 3rd party in order to authenticate with Signals.

Note: Make sure you have pop-up blocking disabled when connecting with a 3rd Party Data Connector

---

- *salesforce*
- *Foresee*
- *Surveymonkey*
- *Gmail*
- *Livechat*
- *Intercom*
- *JIRA*
- *UserVoice*
- *Zendesk*
- *Google Analytics*
- *Trello*

2.2.1 salesforce

After logging in, you'll need to select which report and the field mappings you would like to analyze. If you don't have any reports, more information about them here: https://help.salesforce.com/articleView?id=000113375&type=3

✔ Data Connector Credentials

✏ Data Connector Settings

Parameters

Reports (Required)

These are the reports found in your Salesforce account - please select one you would like to analzye. More information about reports and creating them found here: https://help.salesforce.com/articleView?id=000113375&type=3

| Activities by Salesperson | Age of Cases Currently Open by Type | toma | Accounts and Contacts |

Search...

Search...

⊕ SUBJECT
e.g.

⊕ ACCOUNT
e.g.

⊕ CONTACT
e.g.

⊕ TITLE
e.g.

⤭

Next

✔ Name and Advanced Tools

2.2.2 Foresee

After logging in, we will provide all the Topics and their respective questions. You'll be able to select which questions you want to run an analysis on.

✔ Data Connector Credentials

∨ Answers Bounce

After logging in, we will query your Gmail inbox and pull all the emails found within your query (see below for a link for query options). We will run an analysis on the emails, who they're from, the date and other metadata. ✕

✔ **Data Connector Credentials**

✏ **Data Connector Settings**

Parameters

Email Search Query

This will result in only email snippets returned from this query. You can find out more advanced query techinques here

Include Spam   ☐This will include emails that are filtered as spam.

Next

✔ **Name and Advanced Tools**

2.2.5 Livechat

The LiveChat connector allows you to select a date range and the type of user you want to analyze chats for: Agent, Visitor, or both.

 After submitting your LiveChat username and API key, found here: https://my.livechatinc.com/agents/api-key, we'll pull down your message, chat and visitor data to be analyzed.

✔ Data Connector Credentials

✏ Data Connector Settings

Parameters

Start Date

We will pull data from the specified date until now.

End Date

We will pull data from the specified start date or the beginning of time until the specified end date

Text Analysis (Required)

Agent and Visitor

Please select which text index you would like.

Next

✔ Name and Advanced Tools

2.2.6 Intercom

2.2.7 JIRA

With the JIRA Data Connector, you can select which projects you want to analyze. You can also specify a ticket type if desired. If nothing is specified, Signals will pull all of your JIRA data.

2.2.8 UserVoice

The UserVoice data connector requires a login and api key to connect to your uservoice data. The permissions associated with the login correspond with what data will be available through the connector.

Users can specify which objects you want to analyze.

2.2.9 Zendesk

2.2.10 Google Analytics

2.2.11 Trello

2.3 Public Data Connectors

- *iOS Store Reviews*
- *Google Play Store Reviews*
- *Home Depot Product Reviews*
- *Lowes Product Reviews*
- *Best Buy Product Reviews*

- *Wal Mart Product Reviews*
- *Etsy Shop Reviews*
- *Etsy Product Reviews*
- *Amazon Product Reviews*
- *Twitter* (with twitter account credentials)
- *Facebook* (with facebook account credentials)
- *Weibo search*
- *Youku*
- *YouTube Comments*
- *Consumer Financial Protection Bureau*
- *Indeed*
- *Consumer Affairs*

2.3.1 iOS Store Reviews

After providing the URL for an iOS Store app of your choice, we will attempt to pull down all the text, user, rating, version and date from the reviews (all versions included). The number of text reviews may not match up with the number shown on the app store. If you have any questions, please email clientsupport@stratifyd.com with the URL.

✎ Data Connector Settings

Parameters

iOS App Store URL
(Required)

> https://itunes.apple.com/us/app/pandora-free-music-radio/id284035177?mt=8

The url of the specific app/book/game/movie from the iOS App Store. Browse selection here: 'https://itunes.apple.com/us/genre/ios/id36?mt=8' EX: 'https://itunes.apple.com/us/app/pandora-free-music-radio/id284035177?mt=8

Ratings

☑ 1 ☑ 2 ☑ 3 ☑ 4 ☑ 5

Specify which ratings you would like to search or none to search them all.

Next

✔ Name and Advanced Tools

2.3.2 Google Play Store Reviews

2.3.3 Home Depot Product Reviews

After providing the URL for a Home Depot item, we will download the text, stars, date, location from the reviews and run an analysis on them. If you have questions or concerns, please email clientsupport@stratifyd.com with the item URL.

✎ Data Connector Settings

Parameters

Home Depot Item URL
(Required)

> http://www.homedepot.com/p/Dremel-Saw-Max-6-0-Amp-Corded-Tool-Kit-with-2-Blades-for-Metal-V

Copy the Home Depot item URL and paste it in the field. We can handle different languages.

Next

✔ Name and Advanced Tools

2.3.4 Lowes Product Reviews

After you submit the URL for a Lowes item of your choice, we will download all the customer reviews and run an analysis on them. If you have an questions or concerns about an item you chose, please email clientsupport@stratifyd.com with the item URL.

✎ Data Connector Settings

Parameters

Lowes Item URL
(Required)

> https://www.lowes.com/pd/Dremel-4000-Series-39-Piece-Variable-Speed-Multipurpose-Rotary-Tool-

After logging in, please search for the user you would like to analyze their tweets. Please be aware, you can only run a search or pull a user every 15 minutes in accordance with Twitter's usage guidelines.

✔ Data Connector Credentials

✏ Data Connector Settings

Parameters

Twitter Username
(Required)

@stratifyd

Stratifyd @StratifydTweets   790 Followers

Mira @MiraStratifyd   Followers

Next

✔ Name and Advanced Tools

2.3.11 Facebook

After submitting the Facebook page (not person) URL of your choice, we will pull down all the comments from that page given your chosen date range.

✔ Data Connector Credentials

✏ Data Connector Settings

Parameters

Page URL (Required)

https://www.facebook.com/Hurricanes/

Select the Facebook page you would like to crawl. Ex: 'https://www.facebook.com/PokerstarsUSA/'

Comment Limit(Required)

50000

Sets max number of comments to try to crawl and analyze. (Max 50,000)

Since

12/01/2016   ✕

Until

12/14/2016   ✕

Comments Only

☑ Ignore a user's self posts.

Next

✔ Name and Advanced Tools

2.3.12 Weibo search

2.3.14 YouTube Comments

It can also be used to develop custom connections and/or schedule uploads.

We offer a node.js wrapper for our SDK. https://www.npmjs.com/package/signals-api

To generate an API key:

1. go to Settings in the left-hand menu from the Signals homepage.

2. In the Settings page, click the button to generate an api key towards the bottom.

3. This API key will be generated and downloaded in JSON format to be included in your application.

After the API key has been generated, you can always download it again from the Settings page.
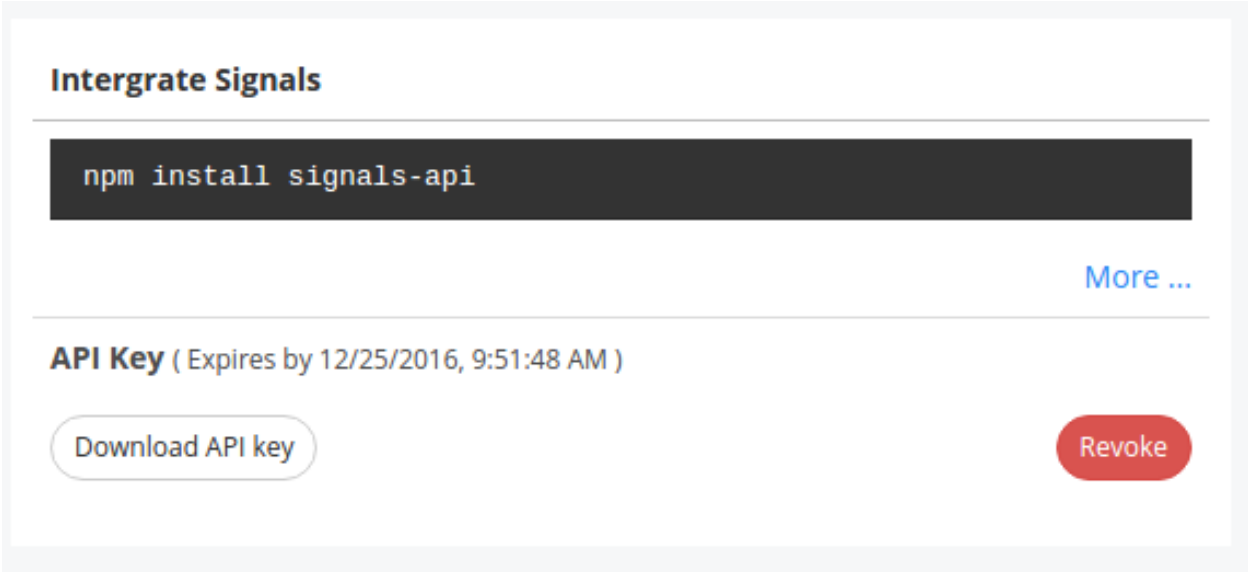
The "Revoke" button will immediately invalidate the API key from further use.

Note: API keys can only be generated by authenticated users. All API keys are tied exclusively to a single user account.

Mapping Data

Signals can be used to analyze any type of data; structured or unstructured.

However, there are 4 types of data that Signals treats differently to give users more information. Those data types are:

1. *Textual* - e.g. product review, customer feedback, news article text, etc.

2. *Temporal* - a date, time, timestamp etc.

3. *Geographical* - location, lat/lon, ip address, etc.

4. *Contributor* - a 1:1 identifier for the creator of the textual data e.g. name, email, GUID, etc.

Data mapping occurs when a new data source is introduced to the platform. CSV, Excel files, `enterprisedataconnectors` and *certain* `3rdpartydataconnectors` require this before analysis.

On the left side of the Map Fields menu, you'll see every field from your data source. Click on the icon to identify any of the 4 data types mentioned earlier.
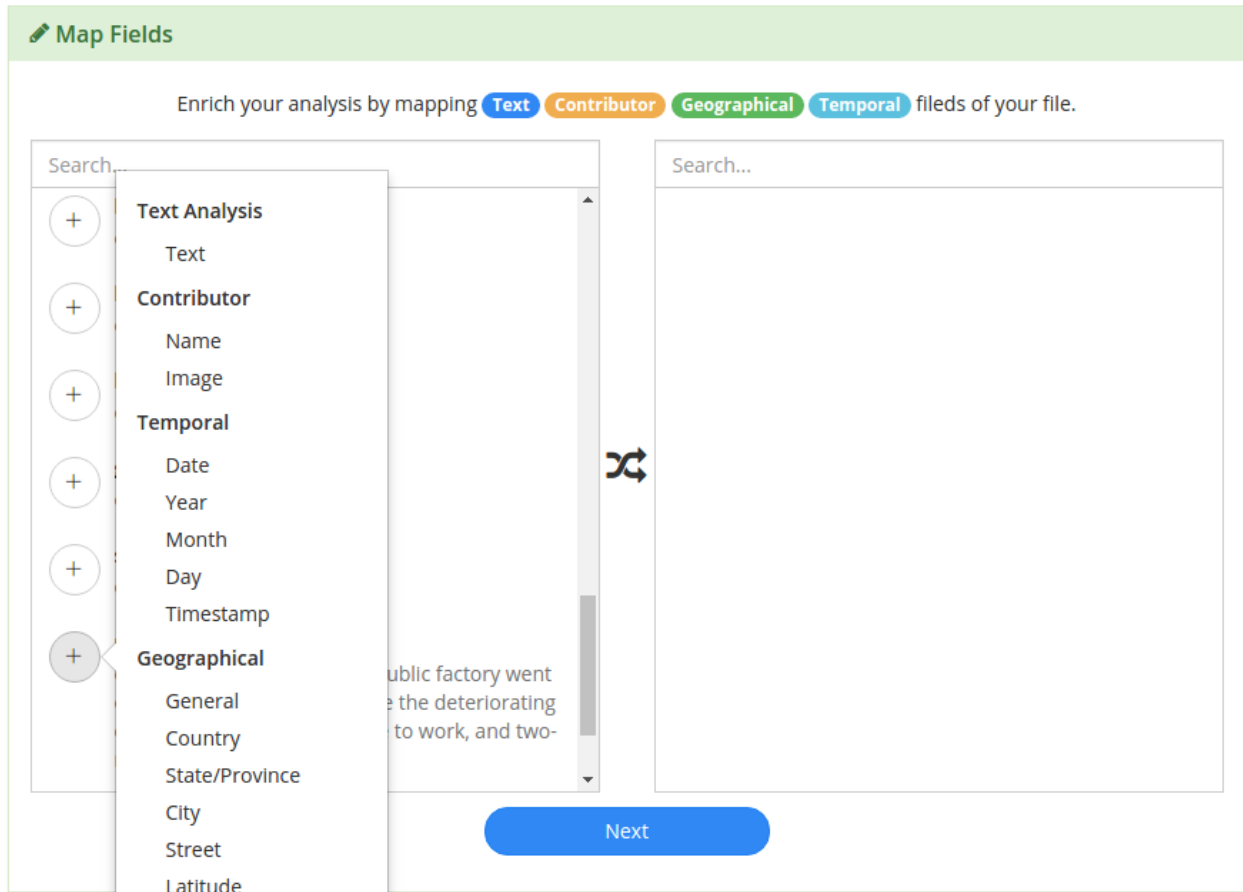
When you've identified the fields, click next and this configuration will be used during processing in the Signals Analytics Engine.
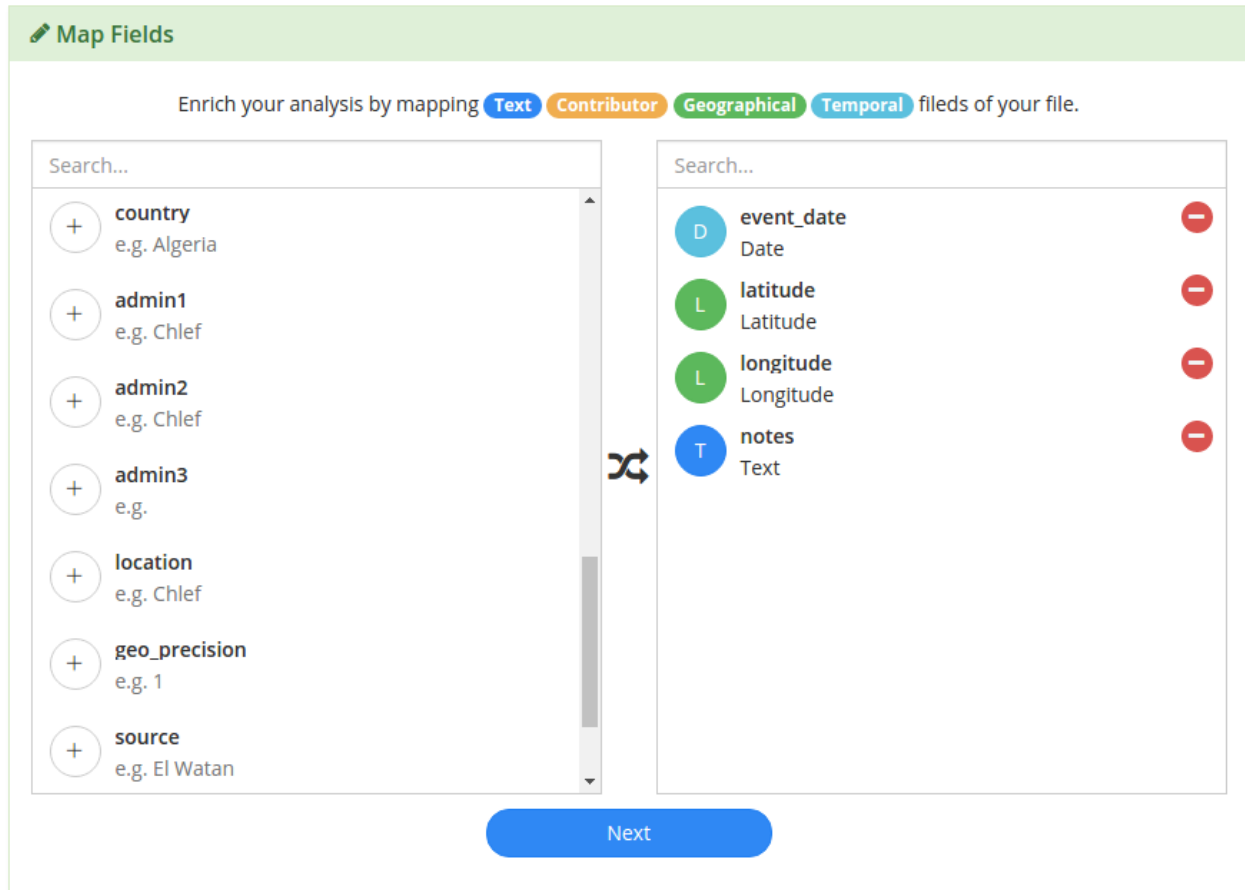
3.1 Textual

The textual data is the data that you want to run unsupervised machine learning to generate buzzwords, a topic model, and sentiment analysis upon. The next section, on *Processing Data*, discusses the outputs of this process.

3.2 Temporal

Identifying where you temporal data is within your dataset allows the Signals Analytics Engine to normalize dates, allowing you the flexibility to roll-up data on different granularities like yearly, monthly, daily, hourly, etc.

Temporal data can be represented in many different formats:

- If your data has a field that contains Day, Month, Year, and/or time in any order, you can simply select the "Date" option in the Map Fields menu.

- If your Day, Month and Year are all contained in separate columns, use the corresponding Map Fields menu items for each to indicate this.

- If your date is represented in some other format, select "Date" and the Signals Analytics Engine will do a best-guess to normalize the data.

3.3 Geographical

The Signals Analytics Engine will create a geographical hierarchy based on the input you provide, giving you access to compare your data at any level, whether it's by country, region, or zip code.

If your data contains multiple geographical or spatial data points, you can indicate where that data resides using the Map Fields menu.

Some of the Geographical options are absolute and therefore don't need additional information:

- For example if your data has a Lat/Long field, Long/Lat field, or one Latitude and one Longitude field, there is no need to specify country, city, or any other geographical information.

- Same goes for IP address and phone number.

- Street, City, State, and Country however, are not 100% by themselves, so map as much information as possible if your geographical data is not absolute.

3.4 Contributor

Contributor data is used to tie all the insights back to an individual to help organizations close the loop with their customers.

This field should represent the person or entity that generated the record or document in your data.

Processing Data

When a dashboard is created with new data, the data goes through the Signals Analytics Engine for processing. This chapter describes what occurs as the engine is processing data. The output is described in the *Analyzing Data* chapter.

Dashboard created, processing your analysis. ✕

0 out of **1** analysis are completed

cfpb 2016-12-14

Performing Geo Tagging
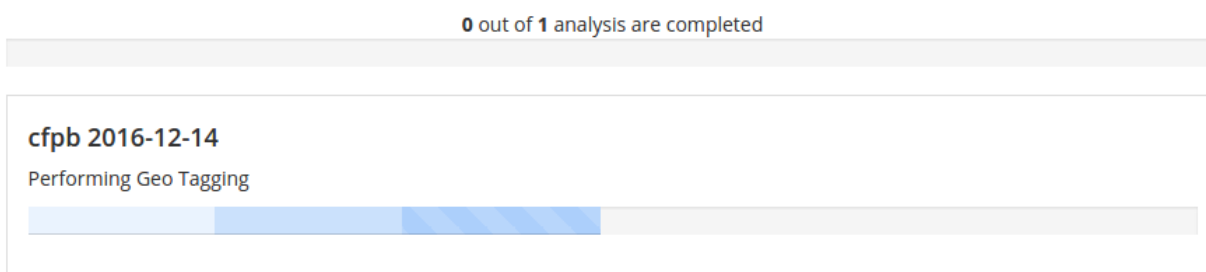
Analysis journey description

Enter Dashboard

The sections below discuss how the output of the *Signals Analytics Engine* can be accessed and used.

Note: To learn how to tune the analytics engine, see the Advanced Options page.

4.1 Signals Analytics Engine

The Signals Analytics Engine leverages Machine Learning and Deep Learning algorithms to help navigate and pivot a large set of textual data.

Built on top of our proprietary Bayesian Neural Network and Generative Model, Signals dynamically identifies semantic topic groups based on the context in your input data.

This is all done in a three-step process:

1. **The engine starts by performing NLP in over 24 languages.**

   In this step, your input documents are *tokenized* into corresponding N-Grams (N>=2), *lemmatized* (words with the same root are grouped together e.g. run & ran), *stemmed*, spam/junk and stop words are filtered

out, and *part-of-speech* tagging and *named entity extraction* are performed. A large N-Gram-based content network is then created based on your input data files.

2. **The engine runs a Multi-Model approach on top of the N-Gram-based content network.**

   This includes using our proprietary text analytics algorithms extended from *Bayesian Neural Network*, *Generative Model*, *LSTM (Long Short Term Memory)*, and *Seq2Seq NLU*. In this step, data input is clustered into semantically meaningful groups.

   The groups are generated and visualized by statistical significance (i.e. the percentage attributed to each topic category in the Semantic Topic Visualization). Each topic is tagged with top representative terms in Buzzwords.

3. **Signals automatically processes all geographical (Where), temporal (When), contributor (Who), as well as any other structured data.**

   It joins the data with the N-Gram-based content network for you to pivot and construct analytics questions against your dataset.

4.2 Reprocessing

There are a number of *Advanced Options* that can be applied to customize the output of the Analytics Engine. When applying advanced options your data will need to be reprocessed. This can either be done directly through Edit Mode, or through clicking the reprocess button on your dataset under Manage Data.
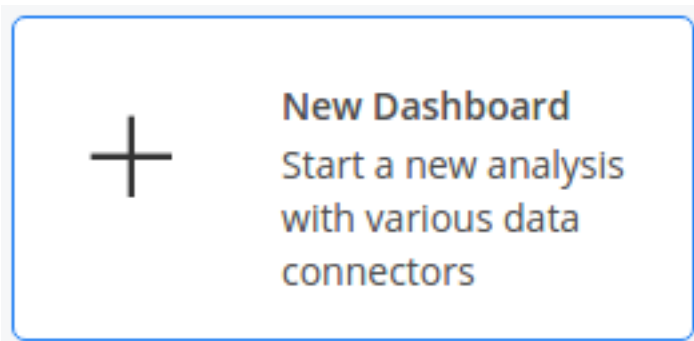
Creating a Dashboard

The dashboard interface is extremely flexible, allowing dashboards to be used for both passive monitoring as well as for active deep-dive or root-cause analysis.

Dashboards are interactive and allow you to visualize and pivot your data on multiple facets, analyzing them holistically or granularly. A Dashboard can hold multiple datasets for disparate data sources.

Dashboards are made up of widgets, or custom visualizations of your data. A dashboard can be split into multiple tabs to create focus areas within your dashboard. The level of connectedness between tabs is customizable. Interactions can affect the entire dashboard, only certain tabs, or just the tab you're on. See the Dashboard Tabs page to learn more about this.

To create a new dashboard, from the homepage, click on the New Dashboard icon:

5.1 Create Your Own Widgets

The widget editor is where you will create the building blocks of your dashboard.

To begin creating a new widget, click on the new widget button:

There are two entry points to begin creating a widget:

1. You can start by choosing the Dimension or Metric you wish to visualize, or

2. You can start by choosing the Visualization type you want to use.

5.1.1 Visualizations

Each visualization supports certain combinations of data types. By selecting a viz first, your available dimensions and metrics will be highlighted.

5.1.2 Options

The options tab will change based on your current visualization.

This is where you can do aggregations, sorting, filtering, and other operations on the data you've selected.

Widget-specific options will also be here, such as the style of map on a geographical widget.

5.1.3 Color

The color tab lets you choose what conditions to color the visualization on.

Select the dimension or metric you want to separate by color, and then choose a method.

Options are:

- Single Color

- Gradient Color (diverging colors based on a range of metrics)

- Palette (pick a color for each dimension)

- Range Palette (pack a color for a range of metrics)
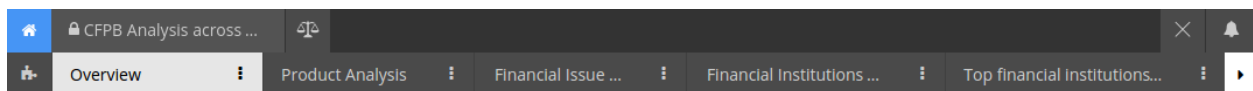

5.1.4 Font

If there is text in your widget, you can specify a custom font in the font tab

5.1.5 Style

The style of an individual widget can be changed here. This will override any customizations on style you've made at the dashboard level.

To see how to change style at the dashboard level, see the Dashboard Styles page.

5.2 Dashboard Tabs

A dashboard can contain as many or as few tabs as desired. Tabs are there to help the designer organize visualizations into logical or intuitive groups based on the analysis context.

To add a tab, simply click the button in the tab menu:

Options can be set at a tab level by hovering over the menu icon on your tab:

Clicking on the edit button will bring up the Tab Detail menu:

By default, the filters from every tab are tied together. Users have the option to set a local query to merge with global queries or to override global queries with a local query. In next chapter, Analyzing Data, this is discussed in detail.

5.3 Dashboard Styles

The aesthetics of any dashboard can be customized. To see options, click on the Customize Style button in the Edit menu:
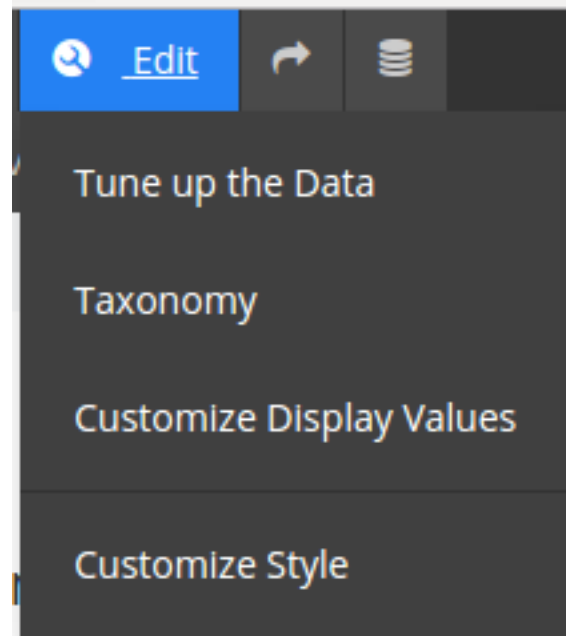
All of the options in the Dashboard Style menu are there to make the dashboard feel like your own:

Dashboard Style ✕

Dashboard Styles	
Background	rgba(255,255,255,0)

Card Styles

Widget Title

⊘ Current

Select

Select

Widget Title

Select

Select

Select

Widget Title

Widget Title

Select

Select

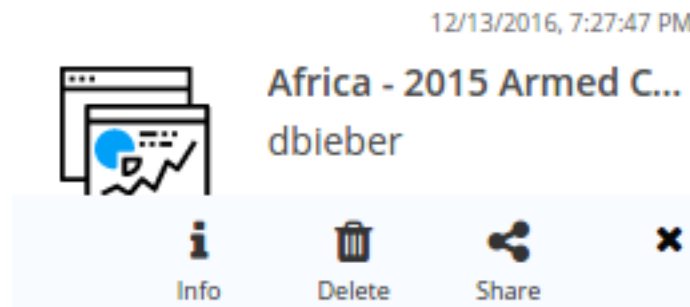**Corner Radius** | 0

Shadow Level | 0

Background | #FFFFFF

Save

5.4 Dashboard Options

Once you've created a dashboard, you can *Permission Levels*, delete, or modify properties by clicking on dashboard menu:

12/13/2016, 7:27:47 PM

Africa - 2015 Armed C...

dbieber

| i Info | 🗑 Delete | ⟨ Share | ✕ |

5.4.1 Dashboard Properties

This is where you can add tags, an image, or even a video, as well as see who else has access.

Note: If a dashboard has been shared with you, you might see a dashboard menu that looks like this:

This means you do not have permission to share with others. For more on permissions, see Administration

---

Dashboard Properties ✕

Name

Africa - 2015 Armed Conflicts

Tags

research ×   public ×

Eg: sales, user experience, social ...

Dashboard Image

| | | Select Image |

☑ Publish to your organization

Share your work within colleagues.

Video URL

Details

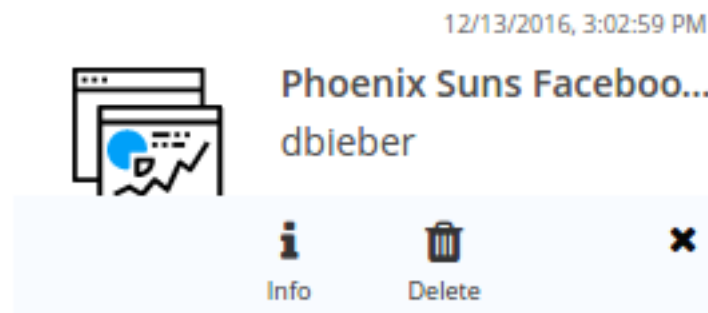👤 Created by **dbieber** at *09/13/2016, 10:37 PM.*
🕐 Last modified by **dbieber** at *12/13/2016, 07:27 PM.*
📢 Your access on this dashboard is **Owner**.
👥 Shared with **dbieber**, **xwang25**, **Public Group**, **st**.
🗄 **1** analysis in dashboard.

Apply

Analyzing Data

This chapter covers how to interpret the data generated by the Signals Analytics Engine as well as how to interact with your data to find meaningful insights.

6.1 Analytics Engine Output

The Signals Analytics Engine generates new datapoints based on your unstructured data. This information is represented in the following way.

6.1.1 Buzzwords

The Signals Analytics Engine reads through every piece of text in a dataset and compares every possible combination of words.

The Buzzwords visualization is made of every statistically significant N-Gram found in the textual data. In order for an N-Gram to be deemed statistically significant, we look at how often they occur together in the dataset vs. how often they occur individually.

6.1.2 Topics

Topics are generated by performing unsupervised machine learning on top of the N-Grams to determine hidden themes and group the documents accordingly.

Any document can occur in more than one topic, therefore the % of documents contained in each topic will add up to > 100%

The Semantic Topics visualization can be represented in a donut chart or a network graph.

6.1.3 Temporal Trends

Temporal information accompanying structured and unstructured data is paramount in understanding quantitative events and their potential underlying relationships across disparate data sets. Signals utilizes time-series predictive analysis, deep learning, and event analysis to uncover trends and patterns across structured and unstructured data.
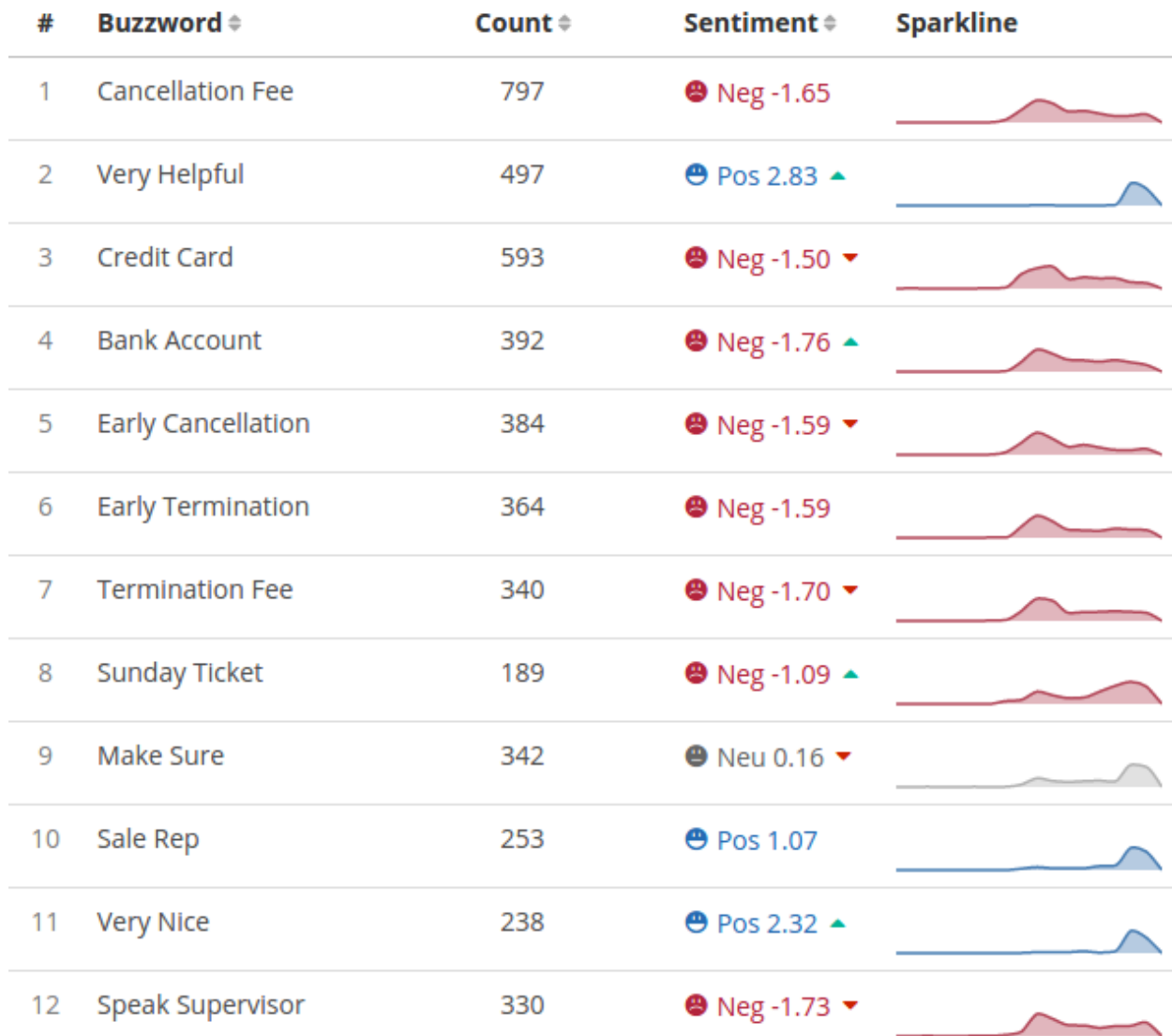
Sunday Ticket Fix Problem Class Action
Debit Card Dish Network Once Again
Take Care Early Termination Lose Signal
Early Cancellation Same Thing Time Warner
Bank Account Sale Rep Loyal Customer Line Sight
Final Bill Credit Card Very Nice Please Help Dish Roof
Feel Like Cancellation Fee Nfl Sunday Best Buy Nfl Ticket
Year Ago Very Helpful Resolve Issue Put Hold
Make Sure Termination Fee Great Job Long Story
Return Equipment Speak Supervisor Nfl Package Six Month
Very Professional Hd Dvr Week Later Very Polite
Past Due Answer Question Very Satisfy
Collection Agency Satellite Dish File Complaint Story Short
Business Practice
Local Channel Look Forward Movie Channel
Gift Card Take Advantage Protection Plan Credit Report
Hold Minute Very Pleasant Two Year Contract Speak Manager
Premium Channel
Question Answer Social Security Attorney General
Better Business Friend Family Overdraft Fee Action Lawsuit

Fig. 6.1: A word-cloud view of N-Grams

#	Buzzword ⇕	Count ⇕	Sentiment ⇕	Sparkline
1	Cancellation Fee	797	😟 Neg -1.65	
2	Very Helpful	497	😃 Pos 2.83 ▲	
3	Credit Card	593	😟 Neg -1.50 ▼	
4	Bank Account	392	😟 Neg -1.76 ▲	
5	Early Cancellation	384	😟 Neg -1.59 ▼	
6	Early Termination	364	😟 Neg -1.59	
7	Termination Fee	340	😟 Neg -1.70 ▼	
8	Sunday Ticket	189	😟 Neg -1.09 ▲	
9	Make Sure	342	😐 Neu 0.16 ▼	
10	Sale Rep	253	😃 Pos 1.07	
11	Very Nice	238	😃 Pos 2.32 ▲	
12	Speak Supervisor	330	😟 Neg -1.73 ▼	

Fig. 6.2: N-Grams in a detail list visualization
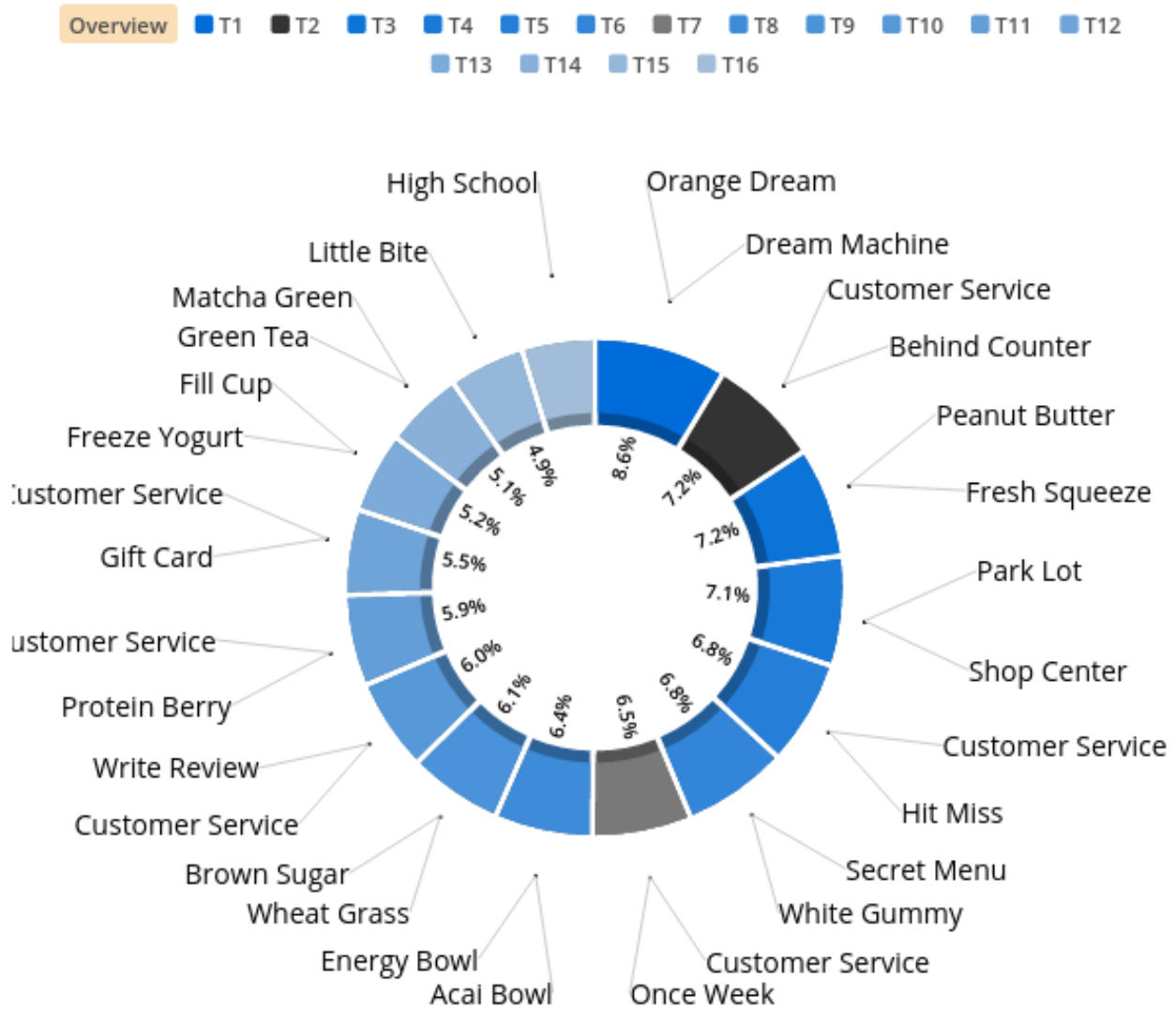
Fig. 6.3: The words around each slice are just the top one or two N-Grams included in that topic. Each topic is represented by a slice of the donut chart. The order (starting at 12:00 and moving clockwise) and size of the slice is determined by the statistical relevance or "tightness" of that topic. The color of the slice represents the sentiment.
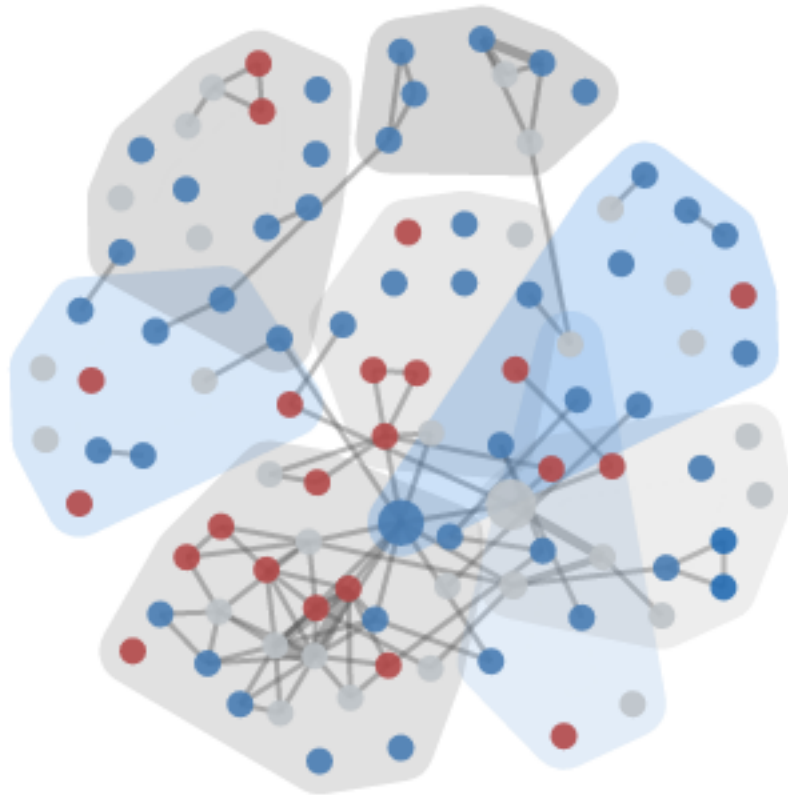
Fig. 6.4: Semantic Topics represented in a network graph of related N-Grams. Each bubble represents an N-Gram, the size indicating the count, and the color indicating the sentiment. The lines show the level of connection, or co-ocurrence, between two N-Grams.
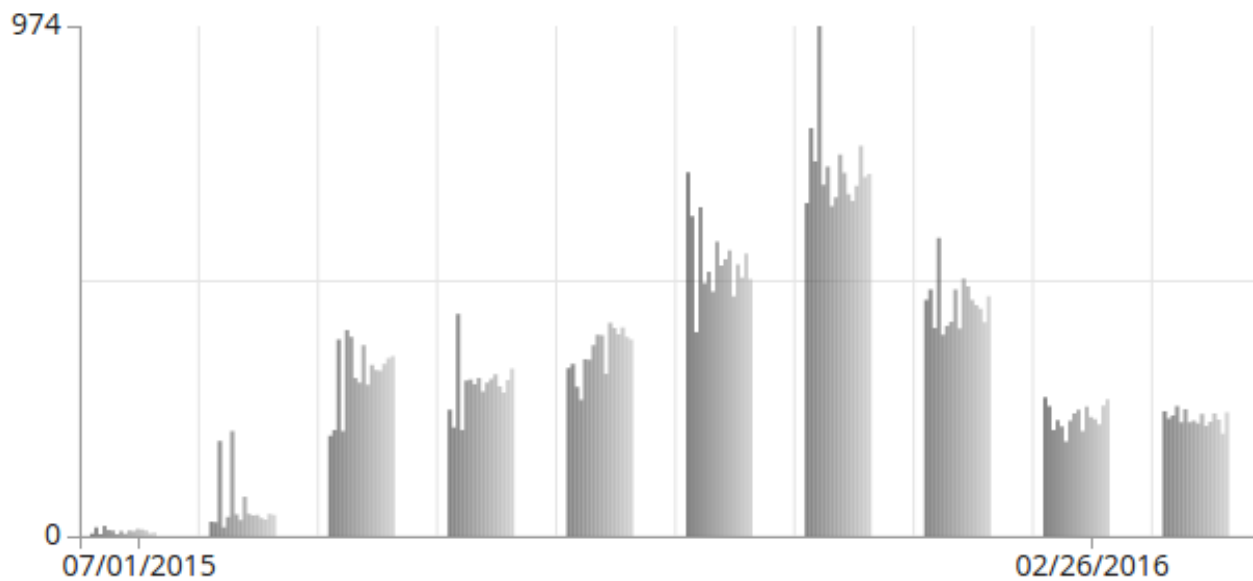
Fig. 6.5: Topic Model trended over time. Each smaller bar represents a topic while the bar groups represent a time period.

6.1.4 Contributors

Contributors help to identify an individual's influence in a data set.

Name: a list of unique identifiers mapped as "name" Count: the # of feedback from an individual or unique ID. Sentiment: The aggregated sentiment score on all the documents contributed by a contributor.

Top Contributors ✏

#	Name ⇕	Count ⇕	Sentiment ⇕
1	👤 steven m.♂	14	😊 Pos 1.14
2	👤 ramsey b.♂	12	😊 Pos 2.00
3	👤 jessica d.♀	9	😊 Pos 2.50
4	👤 angela p.♀	8	😐 Neu 0.77
5	👤 michael c.♂	7	😊 Pos 2.50
6	👤 jennifer c.♀	6	😊 Pos 2.35
7	👤 lisa b.♀	6	😊 Pos 2.50
8	👤 amanda s.♀	7	😊 Pos 1.94

< 1 2 3 4 ... 8 9 10 > Load More ...

6.2 Interactions and Filters

Every interaction in the dashboard essentially applies a filter on your data. This method is designed to help you access insights learned from your unstructured or structured data, and tie it to other dimensions.

When interacting, the filters you've applied will appear at the bottom of the page next to the name of the dataset the filter is applied to.

| citibank | Customer Service | Issue | capital one | Fraud Department | wells fargo | Customer Service | 09/05 12 AM ~ 10/03 12 AM |

Clicking on the name of the filter will remove it.

Filters can also be saved to a dashboard and can be set on one of three levels:

1. *Dashboard*
2. *Tab*
3. *Widget*

6.2.1 Dashboard-Level Filters

To set a filter on the entire dashboard, click the icon in the bottom left corner inside your dashboard. This is the "global" filter panel for your dashboard.

Filters set from here will **not** persist when you leave the dashboard.

6.2.2 Tab-Level Filters

Tab level filters can be applied on a per-tab basis. To do this, click the icon and select one of the two methods:

1. Merge with the current query will allow you to set a permanent filter, but still interact with the data. Interactions with other tabs can still affect the data.

2. Override the current query will apply the filter selections you make and freeze the visualizations so that the tab is no longer interactive. Selections on other tabs will not affect the data.

The main datapoints are available for selection from the Tab Detail page. Other datapoints can be accessed by clicking on the "+" icon next to **Extra Structured Vis Parameters**

Tab Detail ✕

Name

Conflict Details

Filter Override Method

Override the current query ▾

Filters

🔵 ACLED - Armed Conflict Reports: Af...

Topics

T1 T2 T3 T4 T5 T6 T7 T8

Time Range 📅

From 1/1/2015
To 1/1/2016
Sync Selection ☐

Geo +

No region selected

Contributors +

No contributors selected

Buzzwords +

No buzzword selected

Sentiment

Positive Neutral Negative

Extra Structured Vis Parameters +

Submit

6.2.3 Widget-Level Filters

Widget-level filters will override all other filters for that widget.

To apply a widget-level filter, you can either click on the funnel icon in the widget settings menu from your dashboard view, or you can apply a filter from the Create Your Own Widgets in the Options tab.

Fig. 6.6: Widget Settings Menu

Advanced Options

Advanced options allow users to customize the way data is processed. These options are not always needed, but there are some scenarios in which they are very useful. For example:

7.1 Stopwords

Signals provides an out-of-the-box list of stopwords that include the main commonly used non-informative words such as: "the", "an", "I", etc.
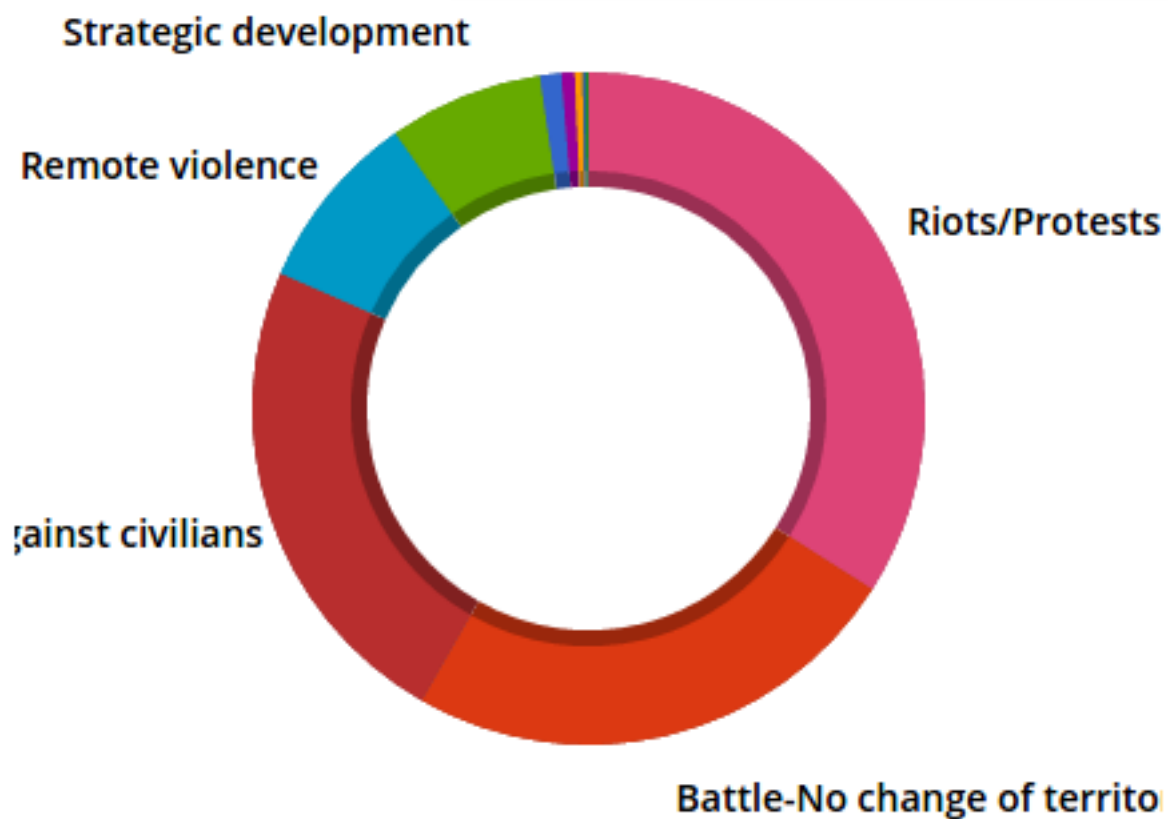
You can create additional lists that can be applied as needed to cut through noise in your textual data on a per-data-source basis.

Typically a brand new data-source is run with the default stopword list, noisy signals can be easily identified in the dashboard and our in-dashboard editor (below) allows you to select N-Grams, topics, contributors, or other features to suppress.

For example, RSS news feeds typically contain the same few sentences at the end of every article:

```
Reporting By Laila Bassam in Aleppo and Tom Perry, John Davison and Lisa Barrington in Beirut;

Writing by Angus McDowall in Beirut, editing by Peter Millership
```

Appears at the end of every news article in some publications. "Reporting By" and "Writing By" will likely be identified as Buzzwords. These terms could potentially link unlrelated documents since they aren't related to the article topics.

To suppress the noise caused by these terms, in the edit menu, click "Tune up data" and select the terms that are non-informative.
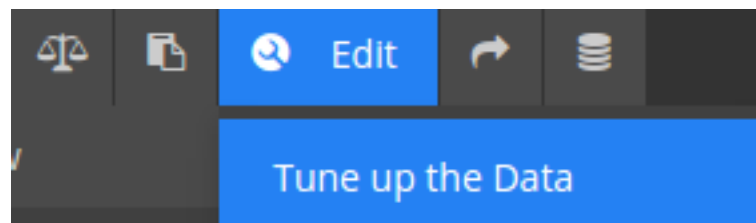
When finished, click submit and your data will begin reprocessing with the feedback you've provided.

7.2 Junk/Spam

Creating a Junk/Spam list can be useful in many different datasets

For example in Twitter data if there is any spam, it will quickly be identified through the visualization. Users can choose to filter out the spam and reprocess the data, indicating to the *Analytics Engine* that the matching documents should be ignored when processing (generating buzzwords and categories).

Note: Spam is typically identified in two ways:

1. A specific user in your dataset is posting junk content - in this case, you can simply select the user to ignore.

2. A specific message has gotten picked up and re-posted by many different users - in this case, you can select the text that is unique to that message and it will be filtered out based on the content.

This is just one example, but as you can see it is generalizable to many other data-types.

7.3 Sentiment

Signals Provides an out-of-the-box sentiment package that is trained on a breadth of data sources however, you have the ability to customize sentiment to tailor results to your specific use case.

Certain terms are generally neutral, but when used in the context of a specific data-set or industry, always have a sentiment associatd with them. Or vise-versa, some terms are very negative generally, but are actually commonly used industry lingo among certain data-sets.

7.4 Taxonomy

Signals supports importing existing taxonomies or you can easily build your own from the ground up. The in-dashboard taxonomy editor allows you to tweak your taxonomy as you analyze your data. Drag and drop buzzwords into your label logic to increase your coverage of existing categories, or as you find new categories.
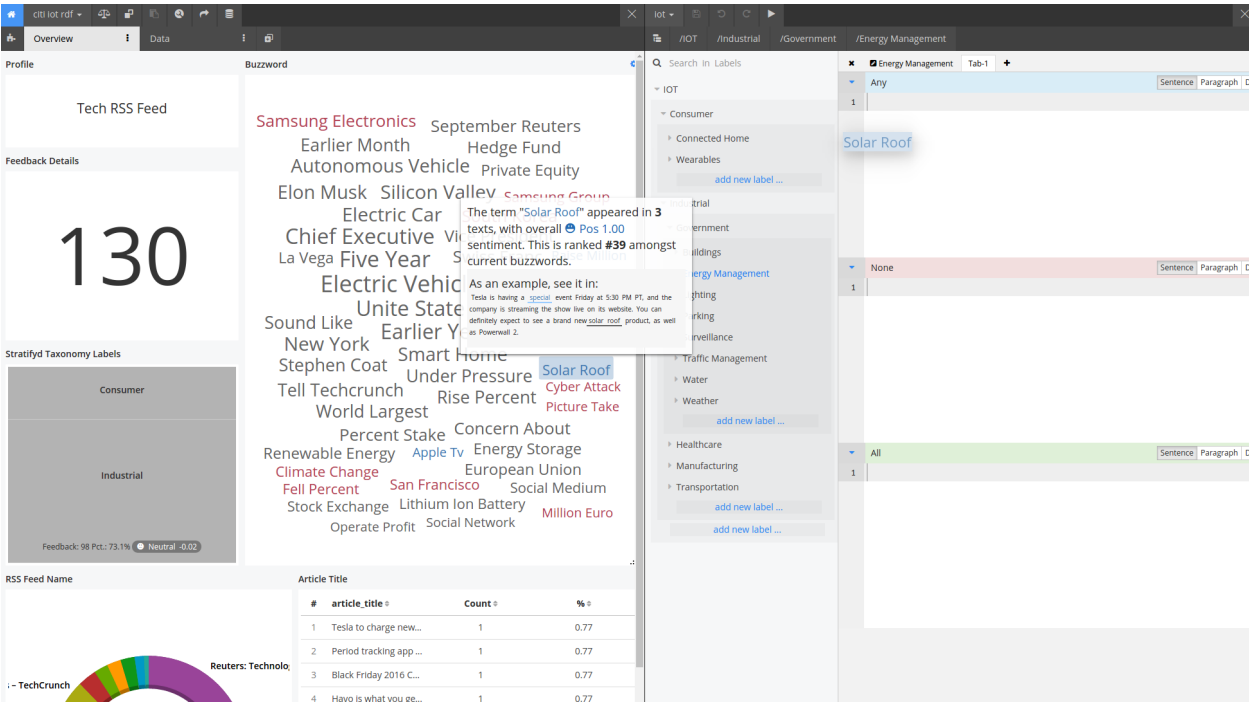
Creating an Account

To register for a trial account on Signals, go to https://signals.stratifyd.com and chat with a support specialist who will help you start your trial period.
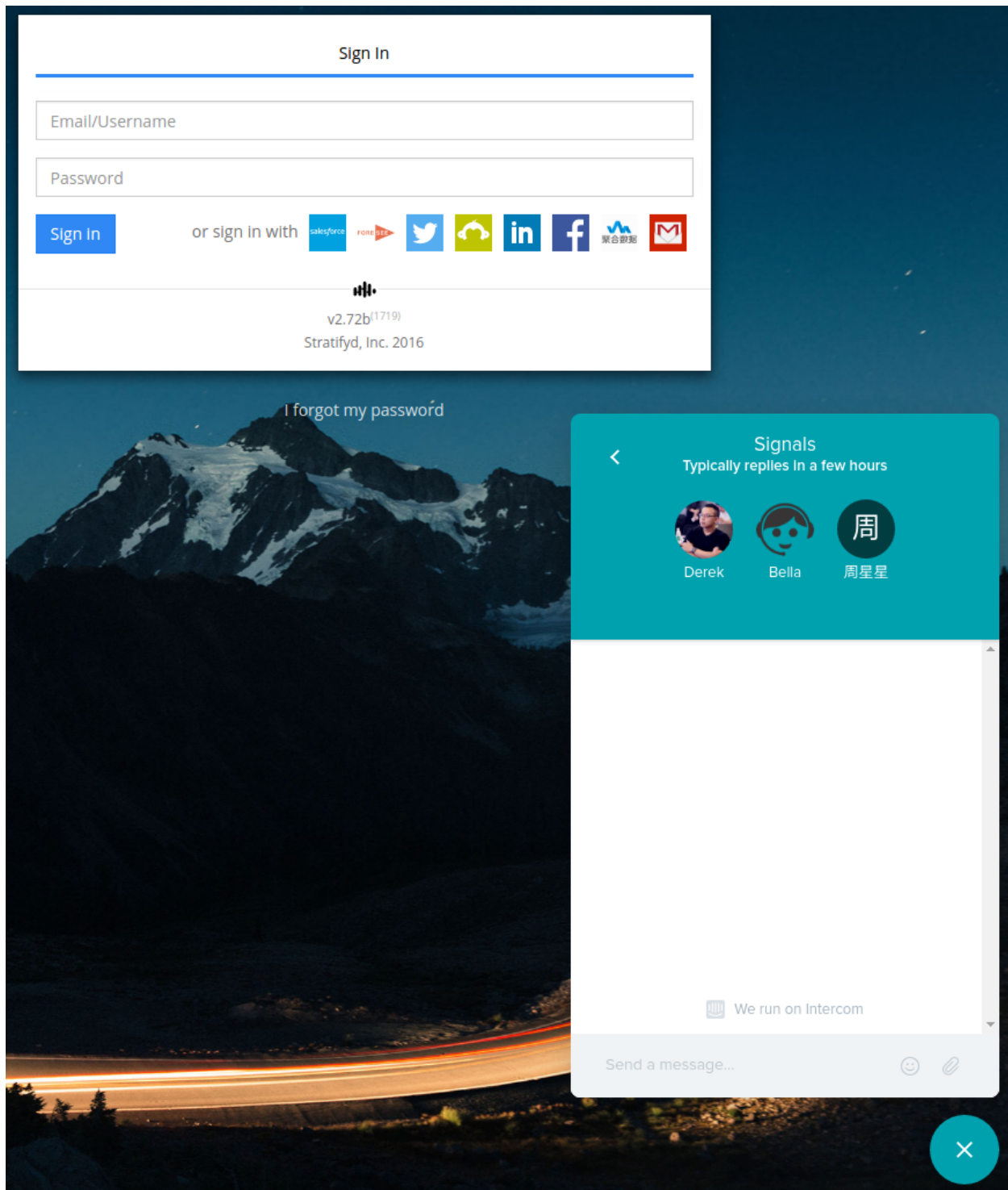
Existing Customers: If your organization has already purchased Signals, you can simply go to

https://[your-organization-name].stratifyd.com

and create an account with your company e-mail and password.

Should you have any questions, please send a note to our Client Support Team or clientsupport@stratifyd.com.

Administration

9.1 Deploying Signals

A **Cloud Deployment** is our most common method of deployment. Among other services, Amazon Web Services S3 and EC2 are used to provide an optimal experience in computing and visualization.

Contact us for information on alternative deployment methods.

9.2 Account Info

Your account info and management console is all available on the Settings page.

9.3 Permission Levels

Permissions on any asset (dashboard, stopword list, taxonomy, etc.) are set by the creator of that object.

Can View users can view and use the asset but cannot permanently modify it.

Can Edit can view and modify the asset, but cannot share it with others.

Can Share can edit and share the asset with others, but cannot remove others' access.

Owner can edit, share, and add or remove users with less privilege.

9.4 User Groups

Group management can be done from the Groups page on the main nav bar.

There are 4 levels of access at a group level:

1. Group Admin
2. Admin
3. Can Edit
4. Can View

The creator of a group is automatically a Group Admin for the group.

Group Admin users:

- have a maximum* access level of **Owner** to any asset shared with the group.
- can add users to the group and give them **Group Admin** rights or lower to the group.
- can remove users with less permission (Admin, Edit, and View Only users)

Note: A group can only have one Group Admin. Group Admins can transfer ownership to another user, but this action demotes the transferer to Admin

---

Admin users:

- have a maximum* access level of **Owner** to any asset shared with the group.
- can add users to the group and give them **Admin** rights or lower to the group.
- can remove users with less permission (Edit and View Only users)

Can Edit users:

- have a maximum* access level of **Can Edit** to any asset shared with the group.
- cannot add or remove other users from the group.

Can View users:

- have a maximum* access level of **Can View** on any asset shared with the group.

*when sharing an asset with a group, the sharer dictates the level of permission granted to the group. The lesser privilege (between permission granted to the asset and permission within the group) will then take precedence for each user. For example if a user grants Admin permissions on a dashboard to a group, group members with Can View in the group will only receive Can View on that object.

Release Notes

Version 2.5 coming early February

FAQ

Where does the name "Stratifyd" come from?

Stratifyd is a unique way of spelling Stratified Sampling, which is a statistical method that illustrates the peeling of layers and layers of data to derive a proper statistical result. We took the name Stratifyd to emphasize next-gen Data Analytics that will be mathematically sound and interpretable by humans. This truly reflects our company mission on providing an Augmented Intelligence environment that leverages Artificial Intelligence to Augment Human's Decision Making process.

For more reading on Stratified: In statistics, stratified sampling is a method of sampling from a population. In statistical surveys, when subpopulations within an overall population vary, it is advantageous to sample each subpopulation (stratum) independently.

Is there a limit to how much data I can upload or analyze?

Trial accounts are limited to 30MB uploads.

While there is no limit on Enterprise accounts, the web interface supports up to 500MB of data. Larger files should be uploaded using a *Data Connector* or via the *Signals SDK*.

How many languages does Signals support?

Signals supports all NLP and Text Analytics functions natively in 25 languages including English, Chinese, Japanese, Spanish, Italian, German, and French.

How do I interpret the word cloud and topic wheel?

Refer to the section in Analyzing Data on *Buzzwords* and *Topics*

Comprehensive FAQ page at http://help.stratifyd.com

Contact

12.1 United States

- 1431 W. Morehead Street Charlotte, NC 28208
- P: 704-215-4955
- webcontact@stratifyd.com

12.2 Silicon Valley

- 4500 Great America Pkwy Santa Clara, CA 95054
- webcontact@stratifyd.com

12.3 China

- 9th, Guanghua road Chaoyang District, Beijing SOHO 2
- P: 1-3466-390016
- paul@tasteanalytics.cn

12.4 London

- 24 Cornwall Rd Dorchester, Dorset DR1 1RX
- P: +44-(0)-7944-724920
- o.bayliss@stratifyd.com

Misc.

13.1 Signals for Students

We are a group of data scientists and PhDs from universities, thus we support students expanding their analytics horizon by providing free access to advanced technologies.

Feel free to drop us a line with your EDU account at clientsupport@stratifyd.com. Our specialists will be more than happy to create an account for you.

If you are a lecturer or professor who wants to use Signals in your classroom, please contact us and we will be happy to arrange licenses for your students. Feel free to drop us a line at clientsupport@stratifyd.com.