

Brainspace

Getting Started



Table of Contents

<i>Log In to Brainspace</i>	3
<i>Open Existing Dataset</i>	3
<i>Data Visualizations Overview</i>	4
<i>Dashboard</i>	5
<i>Cluster Wheel</i>	7
<i>Focus</i>	9
<i>Communications</i>	9
<i>Conversations</i>	11
<i>Thread Analysis</i>	12
<i>Supervised Learning</i>	14
<i>Create Classifier</i>	14
<i>Train Classifier</i>	16
<i>Notebooks, Tags, & Bulk Actions</i>	18
<i>Create Notebooks</i>	18
<i>Create Tags</i>	18
<i>Create Saved Search</i>	19
<i>Bulk Actions & Advanced Search</i>	19
<i>Concept Search</i>	20
<i>Create Concept Search</i>	20
<i>Brain Explorer</i>	22

Log In to Brainspace

To login to Brainspace and access an existing Dataset, follow these instructions:

1. Open a supported web browser and open your Brainspace URL. Please ensure you verify that you have either Google Chrome, Firefox, or Microsoft Edge installed on your computer before logging in to Brainspace for the first time.

Internet Browser	Browser Version
Google Chrome	61 to Current
Firefox	57 to Current
Firefox ESR	52.4 to Current
Edge	41 to Current

Open Existing Dataset

2. From the Datasets dashboard, select which existing **Dataset** you want open. A **Dataset** is equivalent to a project or case in other ediscovery platforms.
3. Each Dataset Tile has a status in the upper right-hand corner. Only **Active** Datasets are accessible.

4

Dataset Dashboard

6

3

5

The screenshot shows the Brainspace Datasets Dashboard. At the top, there is a search bar and a toggle for 'Show only datasets with new documents'. Below this, there are 'Pinned Datasets (0)' and 'Datasets (54)'. The main area displays a grid of dataset tiles. Each tile includes the dataset name, ID, status (Active or Inactive), and total document count. A 'New documents' button is present on each tile. Callout 4 points to the dashboard title. Callout 6 points to the 'Enron_25GB' tile. Callout 3 points to the 'Active' status of the 'Enron_64' tile. Callout 5 points to the 'New documents' button on the 'Enron_64' tile.

Dataset Name	ID	Status	Total Documents
English Wiki	01_6	Active	4,385,721
Enron 1 Million	01_11	Active	1,032,264
Enron 1 Million_64	01_205	Inactive	1,032,235
Enron_25GB	01_204	Active	45,498
Enron_64	01_206	Active	813,186
FastCase	01_279	Active	2,458
GDELT BEC	01_2	Active	2,954,890
George - Reveal Blo...	01_325	Active	600
Hebrew Wiki	01_4	Active	221,574
Hillary Clinton Emails	01_56	Inactive	7,945
ICTS / Protiviti Pilot	01_302	Active	
IRA Russian Troll Tw...	01_117	Inactive	
native data	01_130	Active	
Native Extracted Text	01_207	Inactive	
New DAT dataset	01_312	Inactive	

Last Update: 3/22/2022

4. You can also search for a specific Dataset by name using the search box in the upper left-hand corner of the screen.
5. You have the option of Pinning a Dataset by clicking on the pin in the lower right-hand corner of the Dataset Tile. This will force the Dataset to display within the Pinned Datasets section of the Dataset dashboard for quicker access.
6. Each Dataset Tile provides you with the Dataset Name, Total Documents contained within the Dataset, and then number of new documents that have recently been added to the Dataset.
7. Click on the Dataset to open.
8. Brainspace will direct you to the **Dashboard** data visualization.

Data Visualizations Overview

Brainspace has 5 interactive data visualizations which can be used to analyze data for investigations, compliance events, and ediscovery matters. The section below provides an overview of each visualization and the features that can be used to analyze data.

- **Dashboard** – This visual is used to filter data based on timeframes, term patterns, and facets. Using the Dashboard to filter your data can help you get to the more relevant content faster and help remove low-value data from further analysis.
- **Cluster Wheel** – This visual is used to understand the topics and themes that exist within your Dataset. The Cluster Wheel is the machine's attempt at organizing your data into clusters based on content similarity. Electronic communications and documents that contain similar information or share the same terms/vocabulary, will be grouped together. Each cluster includes a set of themes that provide an overall description of the documents grouped together.
- **Communications** – This visual provides you with an at-a-glance understanding of the communication patterns within your Dataset. It shows which individuals are communicating, how often they communicate, and combine with a keyword or concept search, can show you which individuals are communicating about specific topics.
- **Conversations** – In this visual you can select specific individuals and see how a conversation about a specific topic unfolds over time.
- **Thread Analysis** – This visual provides a graphical representation of email threads including all forwards, replies, and participants within a specific thread.

Dashboard

The graphic below shows each major feature of the Dashboard and how it can be used to filter your Dataset. You can use any combination of the features below to filter your Dataset.

1. **Candy Bar** – allows you to filter out Exact Duplicates, Near Duplicates, and Documents that were Not Analyzed by Brainspace. Not Analyzed documents refer to documents that have no text, too little text, or garbled text.
2. **Document Count** – Represents the Total Documents in real-time. As you apply filters or run searches, this document count will automatically update.
3. **Timeline Graph** – Shows the volume of data across the time period included in your Dataset. You can specify a date field to use for the graph. Email Sent Date is commonly used.
4. **Term Heat Map** - organizes the data by Date/Time and by Concept or Term. The graph shows when there is a high concentration of documents that include a specific term for a specific period of time. The blue shaded cells indicate there a higher concentration of documents related to a specific term for a specific time period; The brighter the blue, the more documents there are for that time period and term. **Anomaly Detection** – Graphs the data based on dramatic fluctuations in the frequency with which the term exists for a specific time period.
5. **Facets** – A dynamic list of values for a specific Metadata field.
6. **Facet List** – There are three (3) Facet Lists on the Dashboard. The first Facet List is the list of Top Terms within your Dataset. This list will automatically update each time you apply a search or filter to the Dataset to reflect the Top Terms within that subset of data. The other two (2) Facet Lists can be customized based on any available Metadata field ingested into the current Dataset. You can click **Load More** to get additional values to display.
7. **Document List** – The Document List provides you with a summary tile like view of each document returned in your filter or search results.
8. **Relevancy Distribution** – This visual is activated when you perform a keyword or concept search. Documents are returned in order of relevance based on a keyword or concept search.

Dashboard

The screenshot shows the Brainspace dashboard interface with several key components and numbered callouts:

- 1:** Points to the top navigation bar containing 'Analytics', 'Notebooks', and 'Supervised Learning'.
- 2:** Points to the search bar area, which includes 'Concept Search' and 'Enron 1 Million'.
- 3:** Points to the 'ORIGINALS' bar in the document volume visualization.
- 4:** Points to the 'Document Volume' heatmap, which shows document activity across months for concepts like 'merlin', 'big horn', 'usam', 'condor', and 'bonus resources'.
- 5:** Points to the 'Top Terms' list on the left side of the dashboard.
- 6:** Points to the 'EmailFrom' list on the right side of the dashboard.
- 7:** Points to the list of search results on the right, which includes 'EXACT DUPLICATE' and 'ORIGINAL' entries.
- 8:** Points to the 'RELEVANCY DISTRIBUTION' graph at the top right of the results panel.

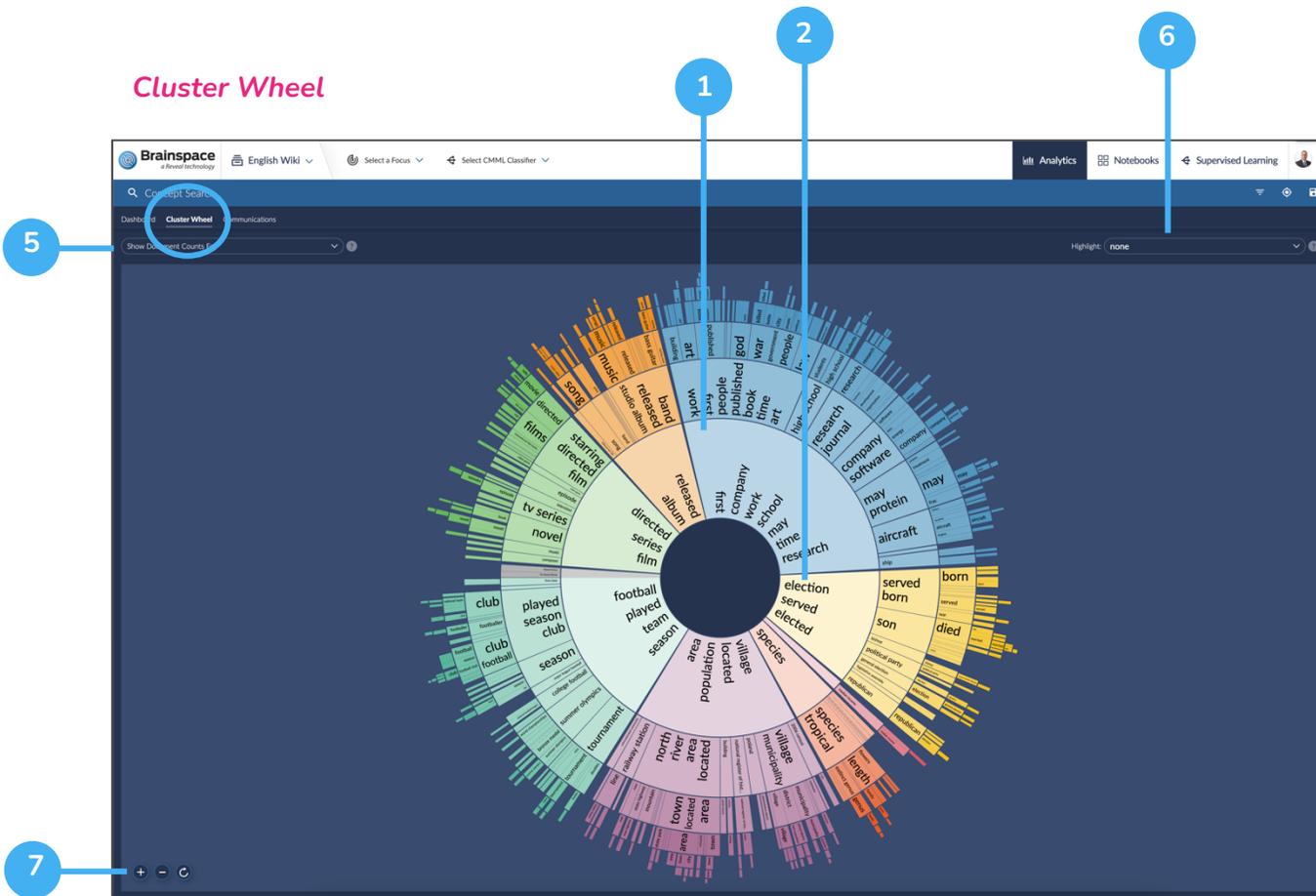
Cluster Wheel

The graphic below shows each major feature of the Cluster Wheel and how it can be used to explore themes or topics within your Dataset. You can use any combination of the features below to zero in on the content you're looking for.

Documents are represented as a circle that surrounds the entire Cluster Wheel. The broader themed clusters are located around the center of the wheel whereas the most specific themed clusters can be found on the outermost edge of the wheel. Broader clusters are broken down into smaller cluster with more specific themes as you move away from the center toward the outer edge of the wheel.

1. **Clusters** – each segment of the Cluster Wheel represents a subset of documents that are associated with a specific Cluster Terms (topics or themes). Click on a cluster to view the Cluster Terms for that cluster of documents.
2. **Cluster Terms** – the topics or themes the best represent the documents contained within a cluster.
3. **Cluster Details Panel** – this graphic shows detailed information about a specific cluster including Cluster ID, Cluster Terms, and Tagged Document Counts.
4. **Cluster ID** – The unique identifier for a specific cluster.
5. **Show Document Counts For** – this feature will display the document counts for tagged documents within the Cluster Details Panel.
6. **Highlights** – Shows you where tagged documents exist within the Cluster Wheel using light shaded clusters to represent a higher concentration of documents and darker shaded clusters to represent a lower concentration.
7. **Cluster Wheel Navigation** – Use these buttons to zoom in or out on the Cluster Terms and to reset the Cluster Wheel back to its original state.

Cluster Wheel



Focus

Once you have applied a filter, selected a Cluster, or performed a search, you have the option of rebuilding the Cluster Wheel on a subset of data. This is called a Focus. Follow the steps below to create a Focus.

1. Click on the **Focus** icon in the upper right-hand corner of the page within any Analytics screen just below the document count.



2. Enter a name for your Focus and select whether you want it to be Public (visible to all) or Private (visible to only you).
3. Indicate whether you want to include family members, threads, exact dupes, and related documents in your Focus.
4. Click **Save Focus** to save and run the Focus creation process.
5. Click on the **Select a Focus** drop-down menu at the top of the page. Your Focus will be listed first while it is being built. There will be a status next to the name of your Focus. Once the build process finishes running, your Focus will be available to access from within this drop-down menu. You can use the Focus Search box to search for your Focus.

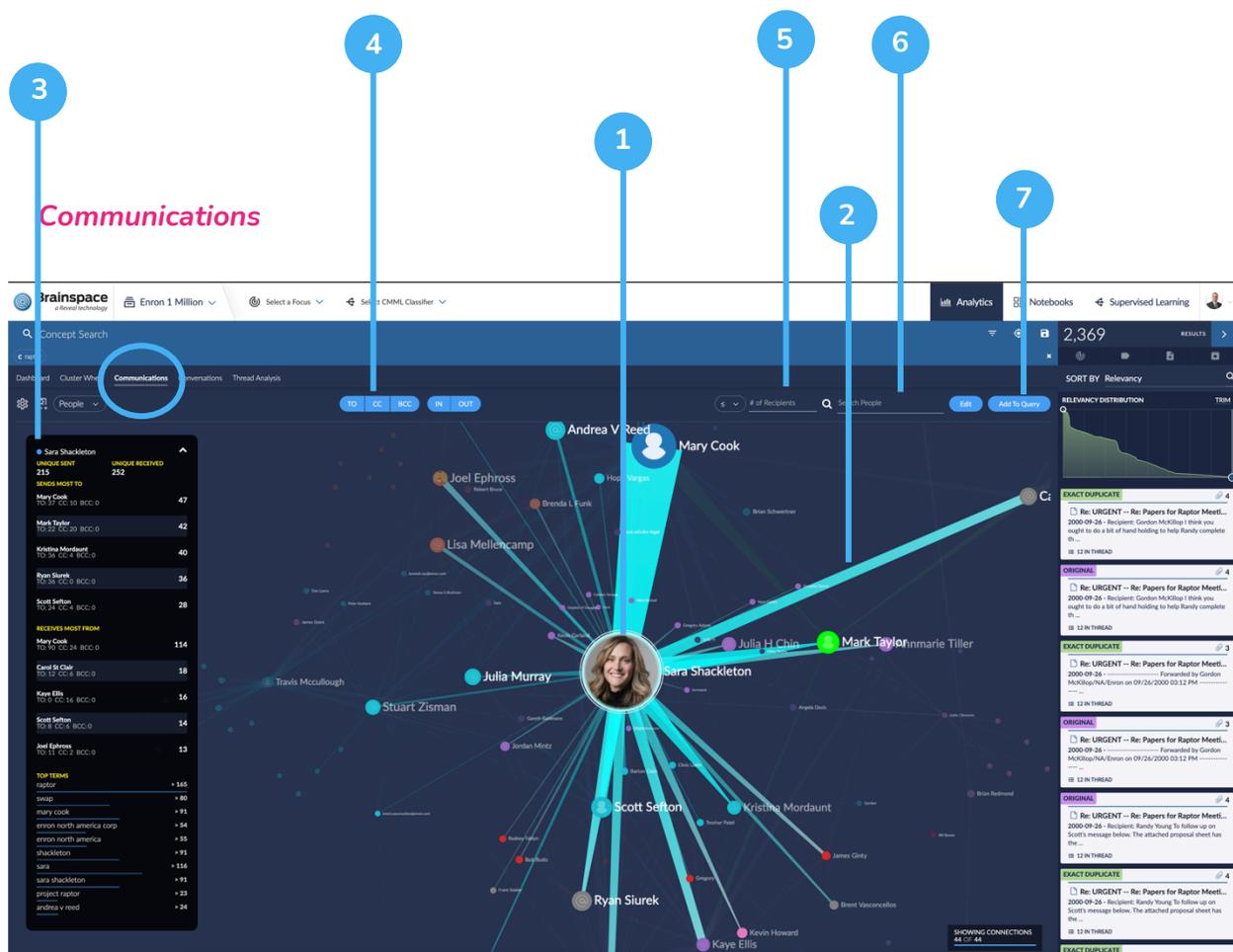
Communications

The graphic below shows each major feature of the Communications visual and how it can be used to understand communication patterns between people and topics.

NOTE: It is recommended that you first apply a filter or run a search prior to accessing the Communication visual since subsets of data make it easier to analyze communication patterns within the visual.

1. **Nodes** – each person or communicating entity is represented as a circle or node within the graph.
2. **Edges** – the lines between communicators. The thickness of the line conveys the volume of communication between Nodes. Thicker lines indicator more communications.

3. **Node Details Panel** – select a Node to see the top individuals the selected person has been communicating with along with the top terms or topics of conversation.
4. **Filter Options** – you can filter the Communications visual so that you only view Sent communications related to a search or filter, communications that include a CC or BCC, communications that were sent to an internal email address, or communications that were sent to an external email address.
5. **Recipient Filter** – You can restrict the visual to a specific number of Recipients. This allows you to view communications that only involve a few people vs emails that go out to an entire department or all employees.
6. **Search People** – You can search for specific people within the Communication visual.
7. **Add Query** – after you have made applied filters, made Node selections, or searched for specific people, you can review the document results by clicking on Add to Query.

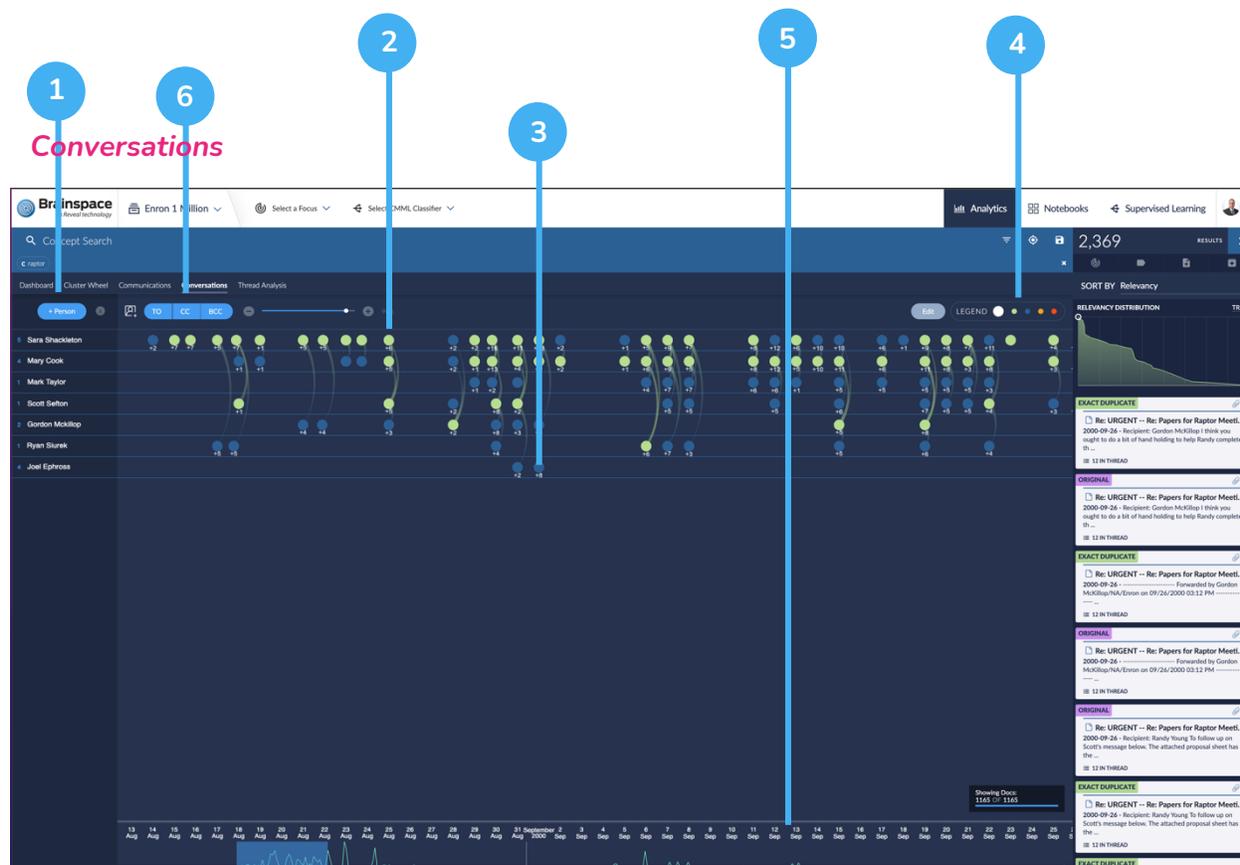


Conversations

The graphic below shows each major feature of the Conversations visual and how it can be used to understand how conversations unfolded over time for a select group of people.

NOTE: It is recommended that you first apply a filter or run a search prior to accessing the Conversation visual since subsets of data make it easier to analyze communication patterns within the visual.

1. **+ Person** – allows you to add people to the Conversation visual. You can add multiple people to see how they communicate over time about certain topics.
2. **Email Nodes** – each node or dot represents one or more emails sent or received to that individual or individuals at a specific time period.
3. **+#** - These are the people that are included in a message but have not yet been added by the user to the visual. You can click the **+#** to add the missing individuals.
4. **Legend** – the colors within the visual and corresponding legend depicts whether the individual Sent, Received, was CC'd, or BCC'd on the communication.
5. **Timeline Graph** – Shows when the communications occurred. Also shows volume of communication across a time period.
6. **Filter Options** - you can filter the visual so that you only view communications received by individuals, communications that include a Cc'd, or communications where the person is Bcc'd.



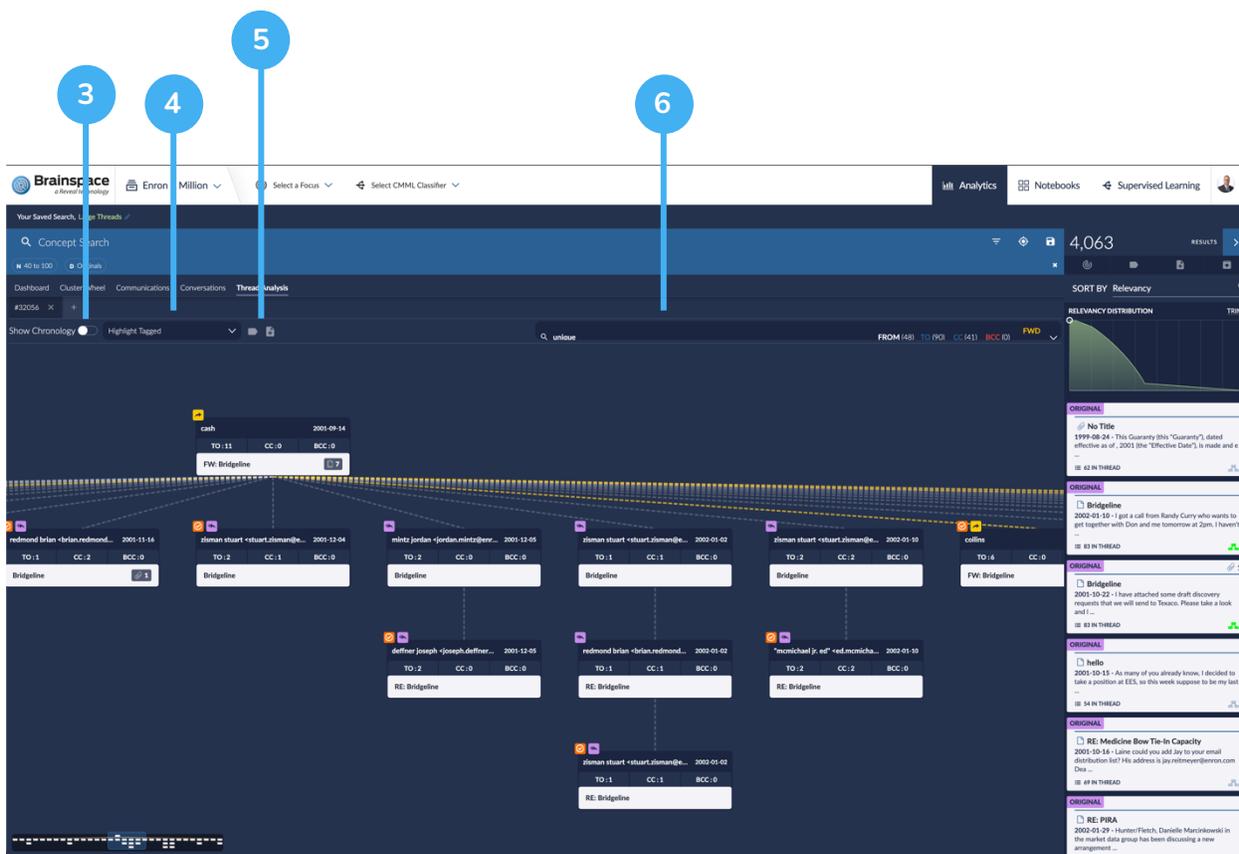
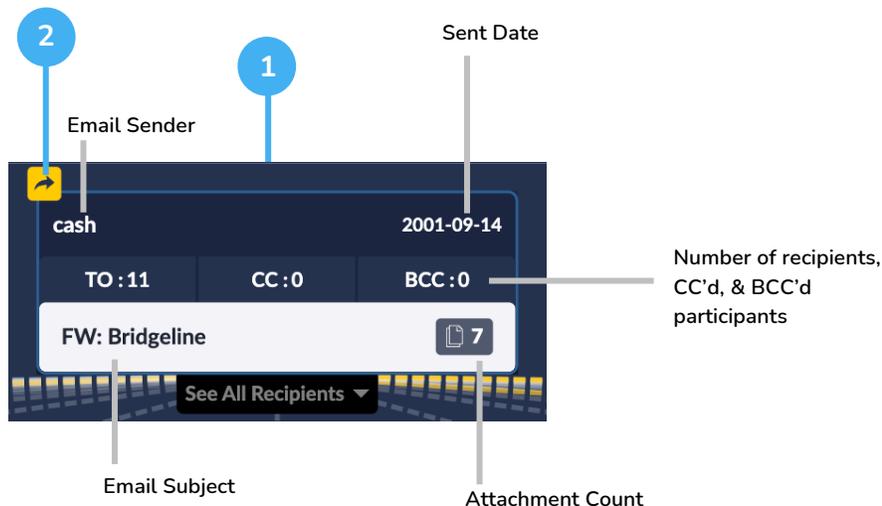
Thread Analysis

The graphic below shows each major feature of the Thread Analysis visual and how it can be used to understand the participants and level of participation associated with email threads.

1. **Email Message Cards** – individual email messages within the email thread.
2. **Direction Flags** – indicates whether the message was a Forward, Reply, or Reply All.
3. **Show Chronology** – displays email messages in chronological order with the thread.

4. **Highlight Tagged** – highlights email messages within the thread that have been tagged with a specific tag and choice.
5. **Bulk Tag and View Options** – allows you to apply a bulk tagged to the email thread and/or retrieve family members, related documents, and exact duplicates.
6. **People Search** – allows you to search for specific thread participants. Also displays the number of forwards, replies, CCs, and BCC emails for participants in email thread.

Thread Analysis



Supervised Learning

Brainspace has a series of supervised machine learning features that enable you to build Classifiers. Classifiers or AI Models, provide you with the ability to train Brainspace to help you find what you're looking for. Using tagged examples, you can train a Classifier to predict whether a document is relevant or not relevant to what you're trying find. These predictions are predictive scores that are assigned to each document within the Dataset. Predictive Scores range between 0 and 1. Scores closer to 1 indicate a document is likely relevant to your investigation or matter based on the manually coded examples. Scores closer to 0 indicate a document is likely not relevant.

Create Classifier

Follow the steps below to create a Classifier or AI Model.

1. Click on the **Supervised Learning** top navigation menu option.
2. Click on **New Classifier** in the upper right-hand corner of the page.
3. Select your Classifier Type which is either **CMML** or **Predictive coding**. CMML is Classifier built using a more TAR 2.0 type workflow whereas Predictive Coding Classifiers follow the more traditional 1.0 workflow.
4. Enter a **name** for your new Classifier.
5. Associate which tag and choice combination should be considered a **Positive** coded example. These are the documents that contain content that are related what you're looking for. To do this, click on **Select a Positive Choice**. Then pick the tag and choice combination. The system will automatically associate the corresponding **Negative** tag and choice combination to your Classifier.
6. Brainspace will display the number of documents tagged for each choice.
7. **Include Metadata** – allows you to include any metadata field content in the Classifier training process.
8. **Choose a Focus** – Limit the documents that are scored to a Focus.
9. **Import portable model** – use the training results of another Classifier to seed you new Classifier.
10. **Enable automatic training** – enable this feature and Brainspace will automatically select documents for you to review and train the Classifier.

New Classifier

The image shows a 'New CMML Classifier' dialog box with the following elements and callouts:

- 4**: Points to the 'Classifier Name' input field.
- 5**: Points to the 'Assign Tag' button in the tag selection menu.
- 7**: Points to the 'Include Metadata' checkbox.
- 8**: Points to the 'Choose a Focus' button.
- 9**: Points to the 'Import portable model...' button.
- 10**: Points to the 'Enable automatic training' checkbox.

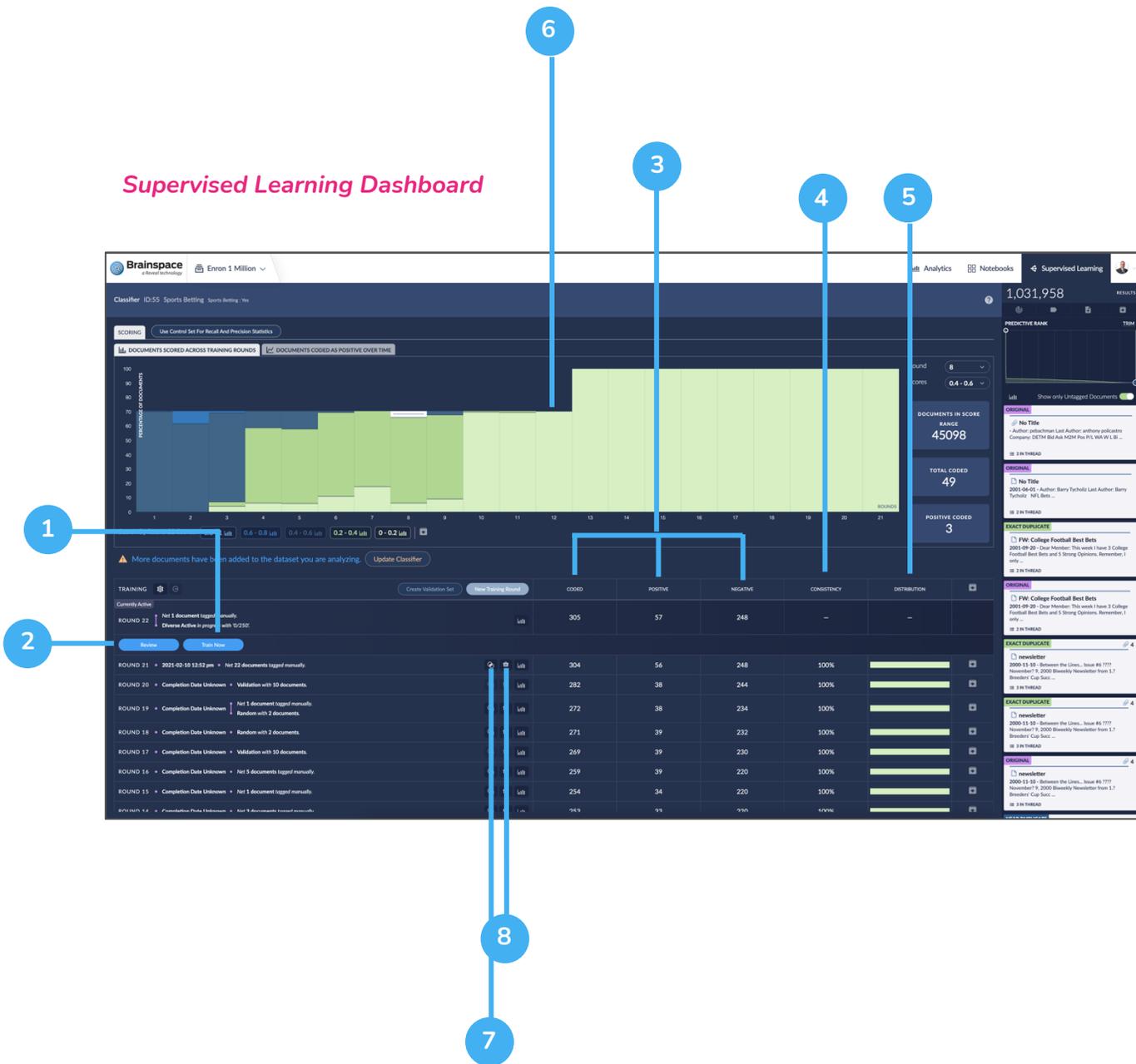
At the bottom of the dialog are 'Cancel' and 'Save' buttons.

Train Classifier

The Supervised Learning Dashboard enables you to train your Classifier and evaluate the results after each training round. Here are the basic features of this dashboard.

1. **Train Now** – allows you to start the Classifier training process for a specific training round
2. **Review** – allows you to review the automatically selected documents that will be used for subsequent training rounds. These are documents selected by Brainspace if you **enabled automatic training** when you created your Classifier.
3. **Round Statistics** – this section provides you with the total documents manually coded along with how many were positive examples and how many were negative examples.
4. **Consistency** – this is the percentage of time that the Classifier predictions agreed with the human coded examples. (how often human and machine agreed).
5. **Distribution** – this bar graph shows the volume of documents and their distribution across the predictive scoring ranges for each training round. Shows how many documents fell into the low, medium, and high scoring ranges for each round.
6. **Documents Scored Across Training Rounds Graph** – shows the the number of documents with the scoring ranges for all training rounds or selected training rounds. Also provides the number of documents, number coded, and number of positive examples within a selected scoring range.
7. **Insights** – allows you to compare scoring fluctuations between rounds for specific terms associated with the Classifier.
8. **Portable Models** – allows you to save the Classifier or Model to the Portable Model Library. You can also download the Model and view the contents. Portable Models can be applied to other similar matter types to accelerate the identification of relevant content.

Supervised Learning Dashboard



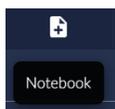
Notebooks, Tags, & Bulk Actions

Create and apply tags, store your work product in Notebooks, and save your searches.

Create Notebooks

Notebooks can be used to store the results of a search, filter, or any subset of documents for further analysis. Follow the steps below to create and use Notebooks.

1. Run a search, filter, or select a subset of data from any visualization.
2. Click on the **Notebook** icon in the upper right-hand corner of the screen within the Analytics tab.



3. Click on **New Notebook** to create a Notebook or select an existing Notebook from the available list.
4. Enter a name for the Notebook
5. You can make the Notebook Public or Private, included related documents, associated specific Tags, or associate specific Classifiers.
6. Click **Create Notebook** to save your new Notebook.
7. The subset of documents displayed will be automatically added to your new Notebook.
8. To access any existing Notebook, click on the Notebooks tab in the top right-hand corner of the screen.

Create Tags

Tags can be used to recall documents that are related to specific issue. You can create tags and use them to categorize your analysis and review. Follow the steps below to create a tag

1. Click on the down arrow in the upper right-hand corner and then click on the Administration menu options.
2. Locate the Dataset you want to create tags for and click on the Tag icon. 
3. Click on **New Tag**
4. Enter a name for your new Tag
5. Create the choices for your tag. Example: Relevant, Not Relevant
6. Click the + button for each additional choice you want to create for your Tag
7. Click Save
8. You can now use this tag at a document level or to bulk tag a set of documents within your Dataset

Create Saved Search

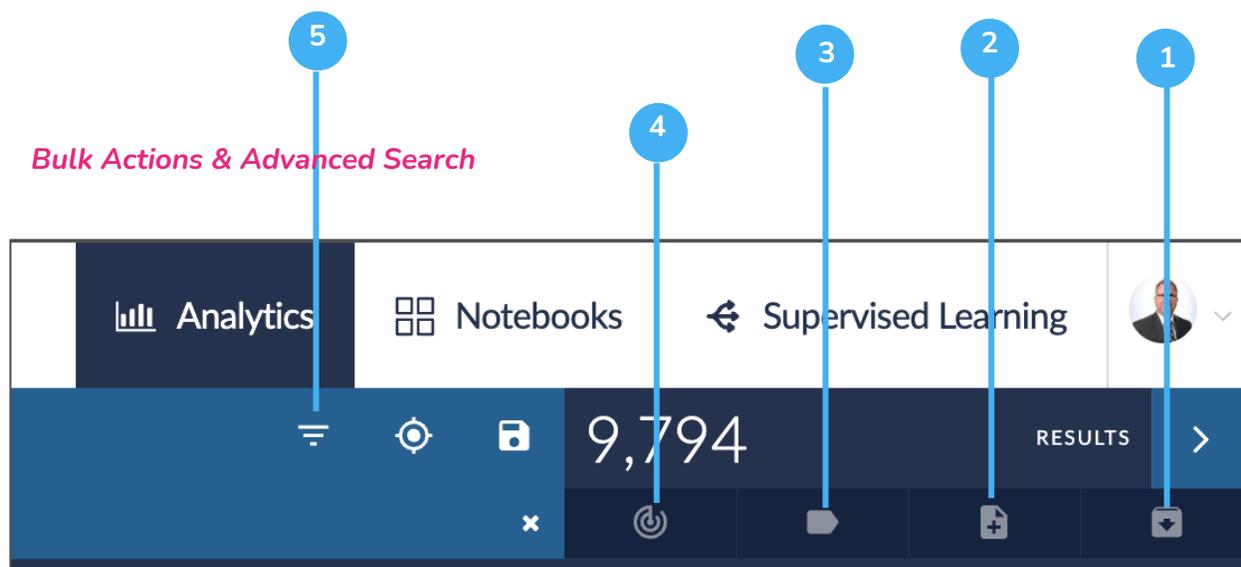
You can build a search using any filter, concept search, or the advanced search options. To save the search you've built, follow the steps below.

1. Click on the **Saved Searches** icon in the upper right-hand corner to the left of the document  count.
2. Click on **Save Current Search**. 
3. You can recall your saved search from the same **Saved Searches** menu option.
4. You can recall your search history by clicking on the **Search History** icon. 

Bulk Actions & Advanced Search

Here is a list of bulk actions that you can perform and where these features are located within the Analytics tab.

1. **Download Search Results Report** – allows you to download a report of the search results including any metadata associated with the documents in your results
2. **Notebooks** – you can bulk add a subset of documents to a Notebook
3. **Tags** – you can bulk tag a subset of documents
4. **Focus** – you can create a Focus on a subset of documents
5. **Advanced Search** – allows you to run keyword searches and metadata filters against the Dataset.



Concept Search

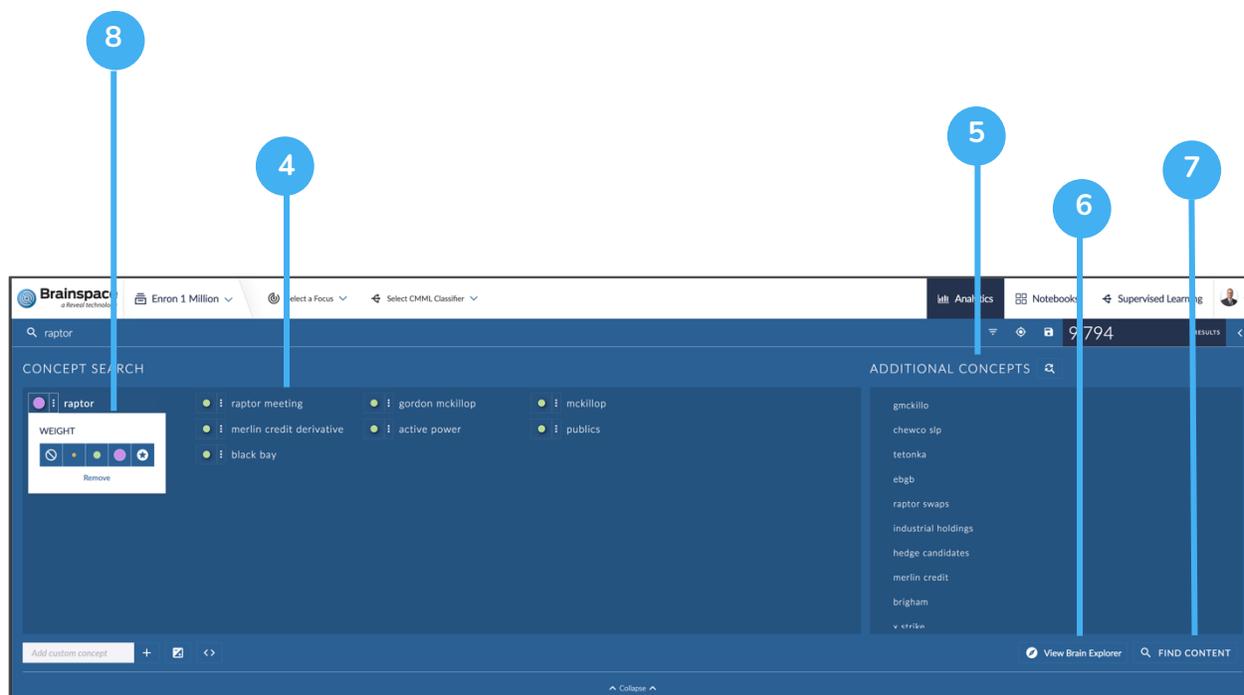
Concept Search can be used to locate relevant content by searching for a single word, phrase, or even a paragraph. As you search for terms, Brainspace will return your search results along with a set of related terms or concepts that are related to your original search. You can explore these related concepts and decide if there are additional terms that would be important to your analysis.

Create Concept Search

Follow the steps below to create a new Concept Search.

1. Enter a term, phrase, or copy and paste a block of text into the **Concept Search** text box. A **Concept Search** can be executed within any visual in the Analytics tab.

2. The search results will be sorted in descending order by a **Relevance Score** which is assigned to each document in the search results based on how often the term appears within the document and how often the related concepts appear. The documents with the highest Relevance Scores appear at the top of the search results.
3. The Related Concepts drawer will also appear after you run your Concept Search. See diagrams below for explanation of each feature of the Related Concepts.
4. **Related Concepts** – list of the top related concepts/terms to your original search concept search.
5. **Additional Concepts** – the second tier of related concepts to your original search.
6. **View Brain Explorer** – launch the Brain Explorer visualization. See Brain Explorer section below for description of feature.
7. **Find Content** – allows you to re-run the Concept Search after you've made additional selections within the Related Concepts drawer.
8. **Concept Weights Setting** – you can adjust the importance weight of a specific concept which will affect the sort order of your search results. You can also make a concept required to appear in ALL documents within your search results.



Brain Explorer

The Brain Explorer is used to visually understand the relationship between your Concept Search and the related concepts. You can use this visual to modify or tune your Concept Search.

1. **Concept Weights** – this legend shows the color assigned to each concept weight category. The Concept Weight identifies the level of importance of the concept or term
2. **Concept Nodes** – each node represents a concept. Click on a node to see other related concepts associated with the selected concept.
3. **Additional Concepts List** – additional related concepts that do not appear in the visual but may be meaningful to the search.
4. **Add Custom Concept** – You can add a custom concept to see how it relates to the existing concepts/terms displayed within the visual.
5. **Additional Concepts Per Node** – decide how many related concept nodes to display for each selected concept.

