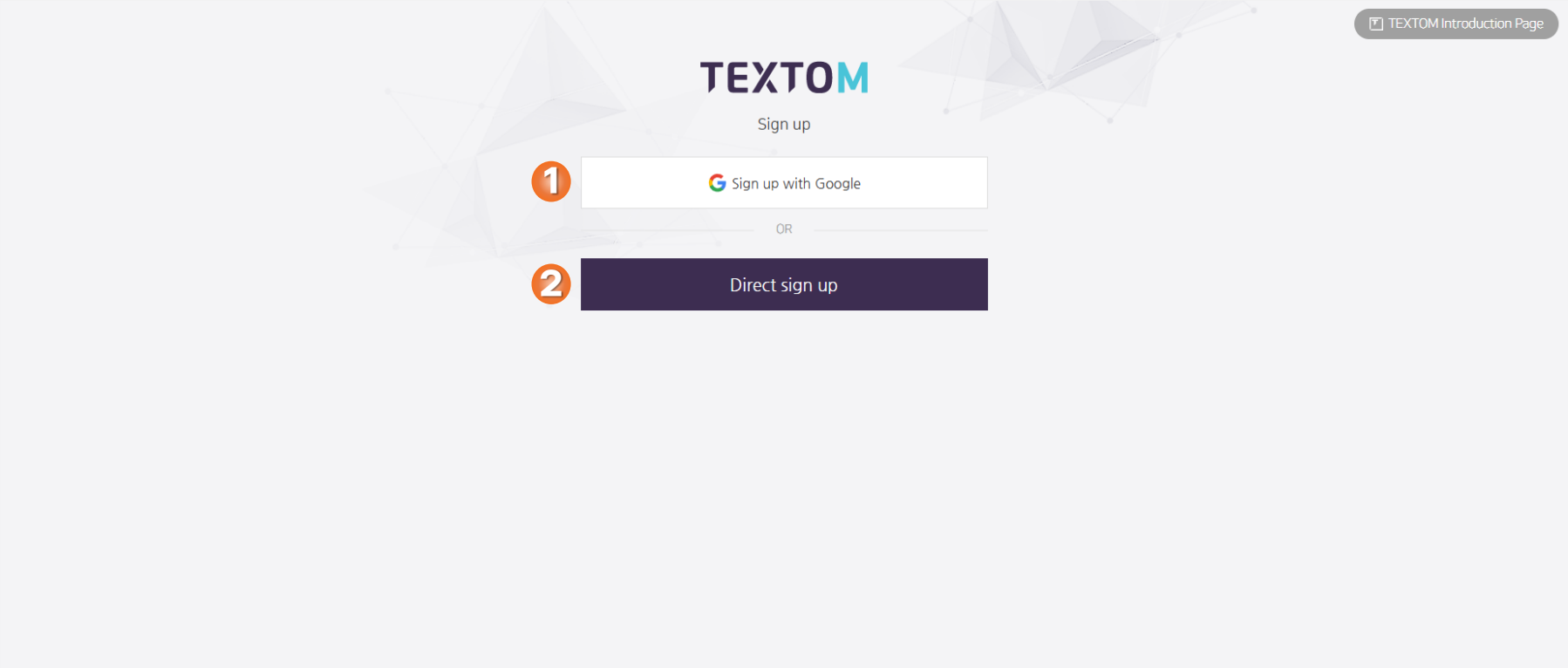

BIG DATA ANALYSIS PLATFORM TEXTOM
TEXTOM
MANUAL
TEXTOM

TEXTOM

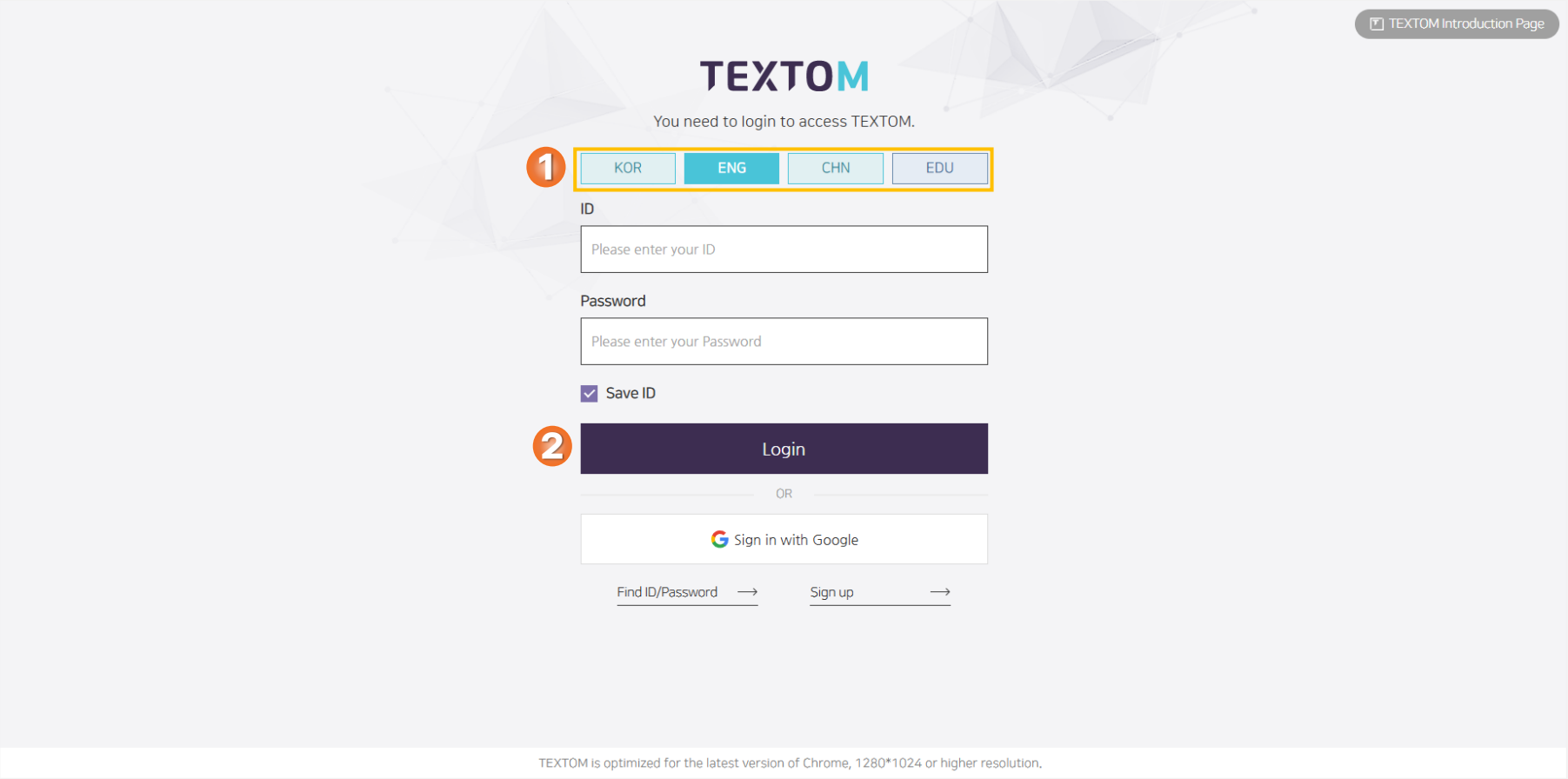
CONTENTS

Click the table of contents to go to the page.
Review the manual before use for easier navigation of TEXTOM



① **Sign up with Google**
Easily sign up using your Google account.

② **Direct Sign up**
Sign up by entering your personal information.



① Select Service

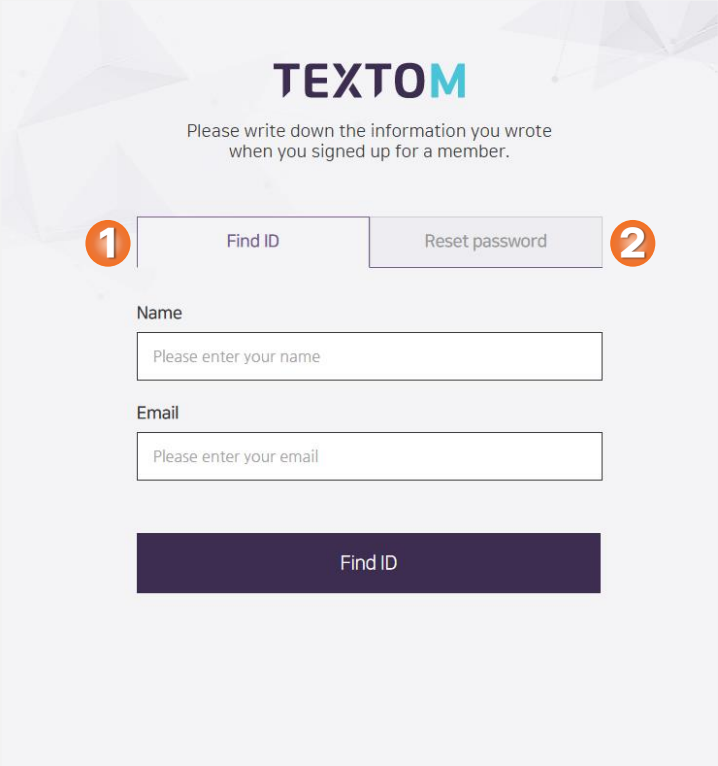
Log in to the desired service : TEXTOM, TEXTOM Global, TEXTOM China or TEXTOM Edu.

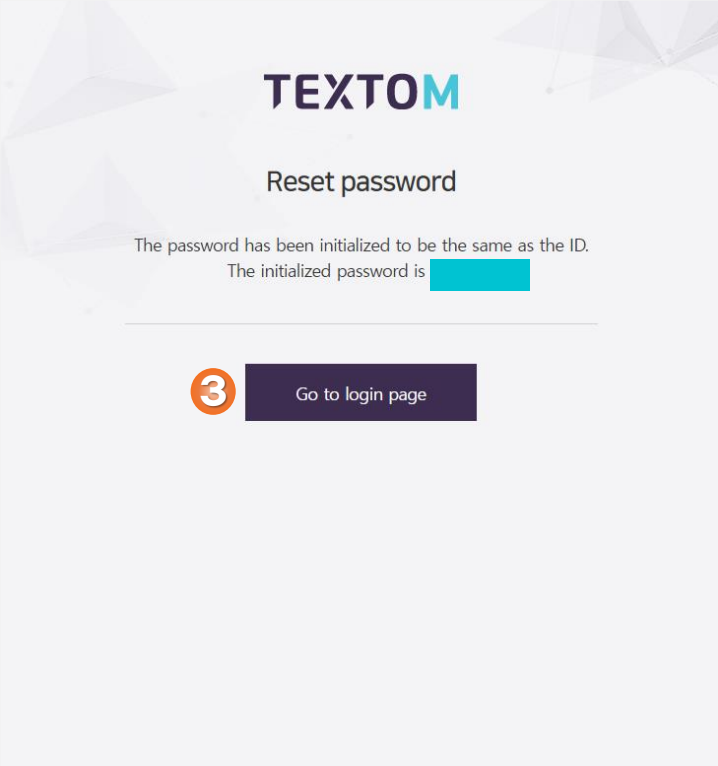
Tip. Existing members can use email authentication to access all three services (Korean, English, and Chinese).

Tip. To log in to TEXTOM Edu, you must first have a TEXTOM account.

② Log in with Google

You can easily Log in using your Google account.





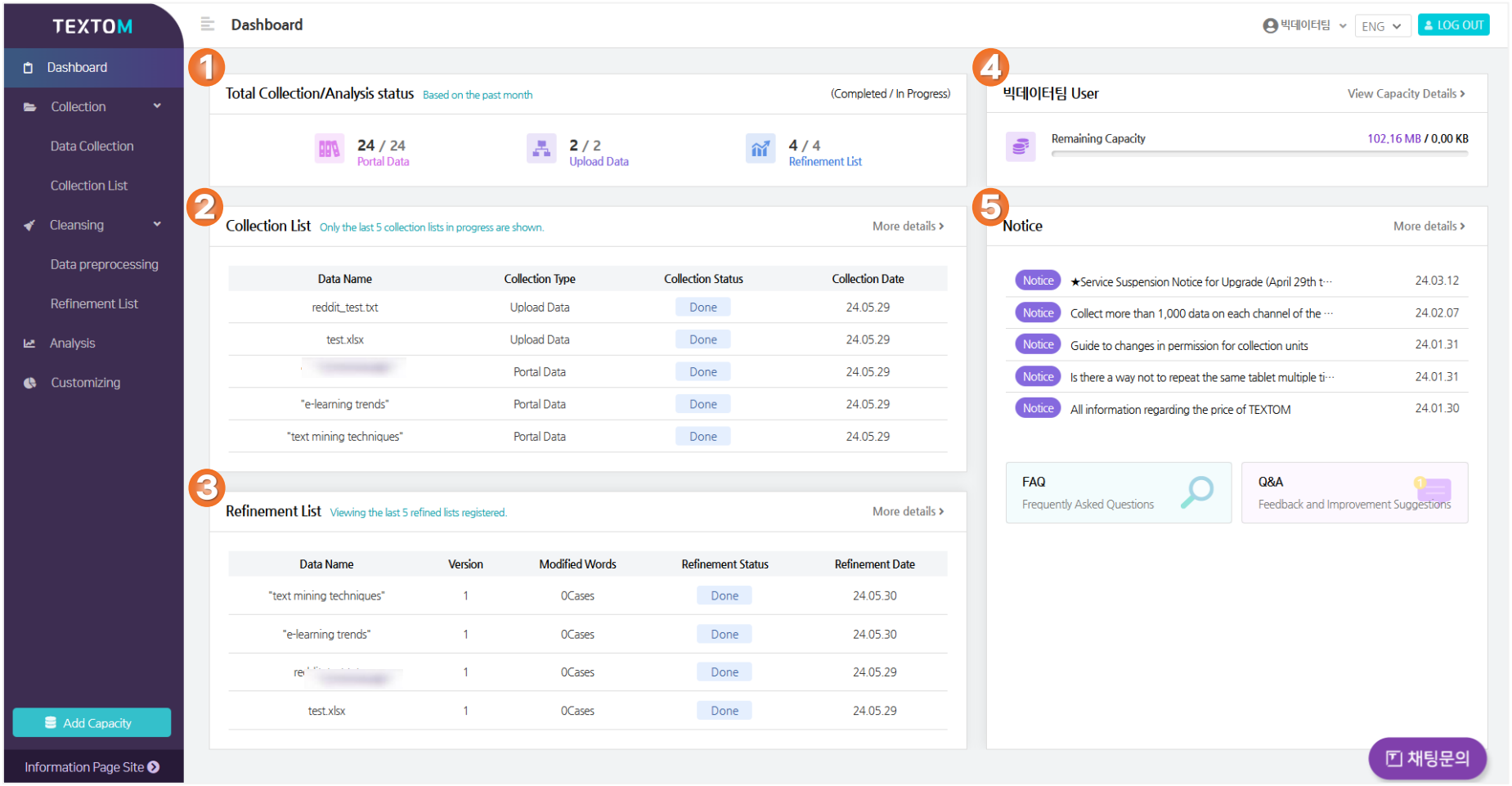
① Find ID
You can find your TEXTOM ID by name and email.

② Reset password
You can reset your existing password by entering your ID, name and email address.

③ Reset new password
You can reset your password to a new one on the [My Page]-Edit Member Information page.
Tip. The reset password will be the same as your current TEXTOM ID.

Dashboards

You can check the overall analysis status.



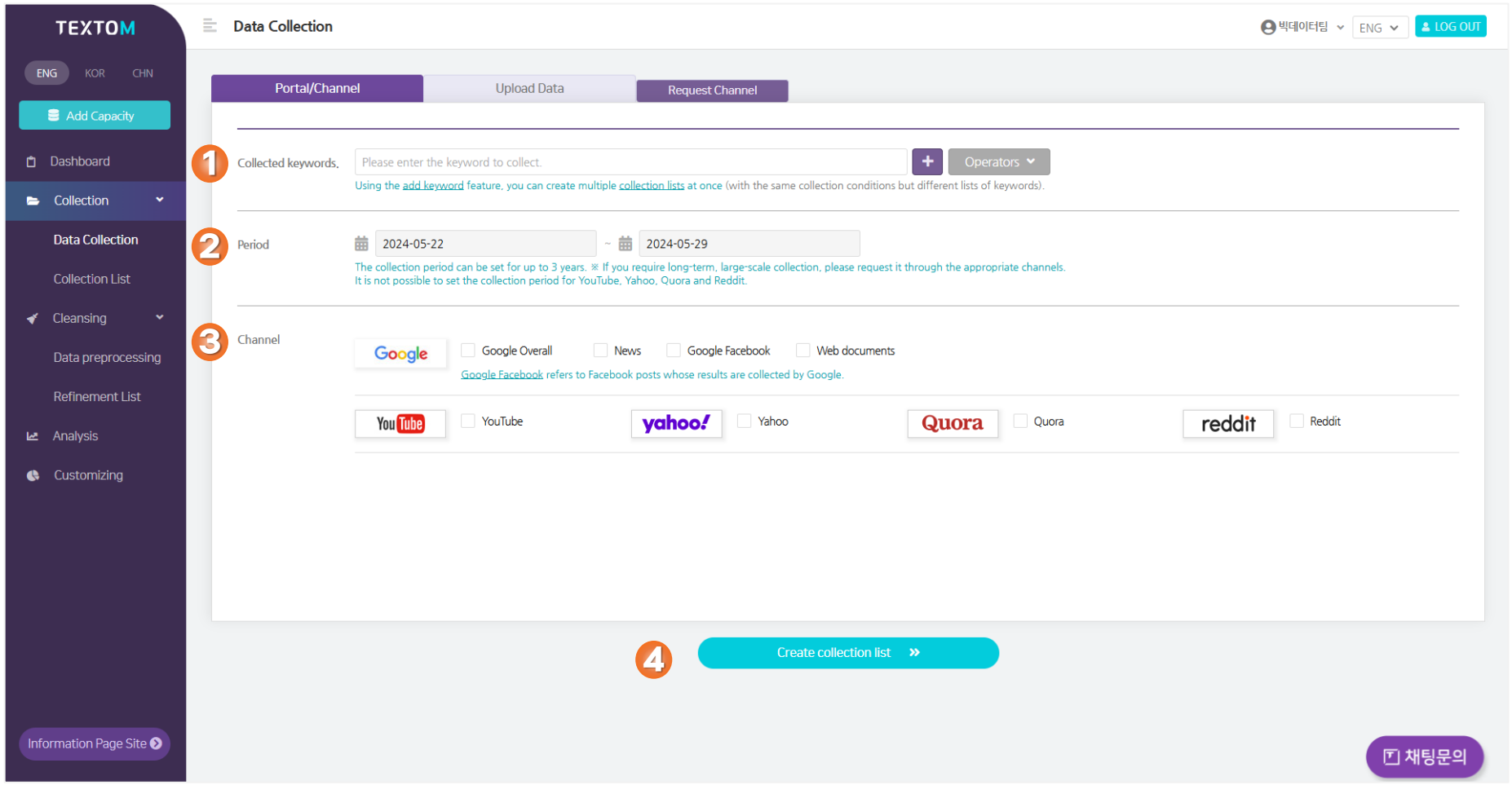
① **Total Collection/Analysis status**
The current status of data collection and refinement can be checked for each channel : Portal, Upload Data.
Tip. Clicking the icon will take you to that page.

② **Collection Progress List**
Check the collection progress of the last five registered data.
Tip. Clicking [More details] button will take you to that page.

③ **Data Preprocessing List**
Check the refinement progress of the last five registered data.
Tip. Clicking [More details] button will take you to that page.

④ **Remaining Capacity**
Check the remaining data.
Tap [Capacity details] to go 'My page'.

⑤ **Notice**
View service information such as announcements, manuals, and utilization papers.

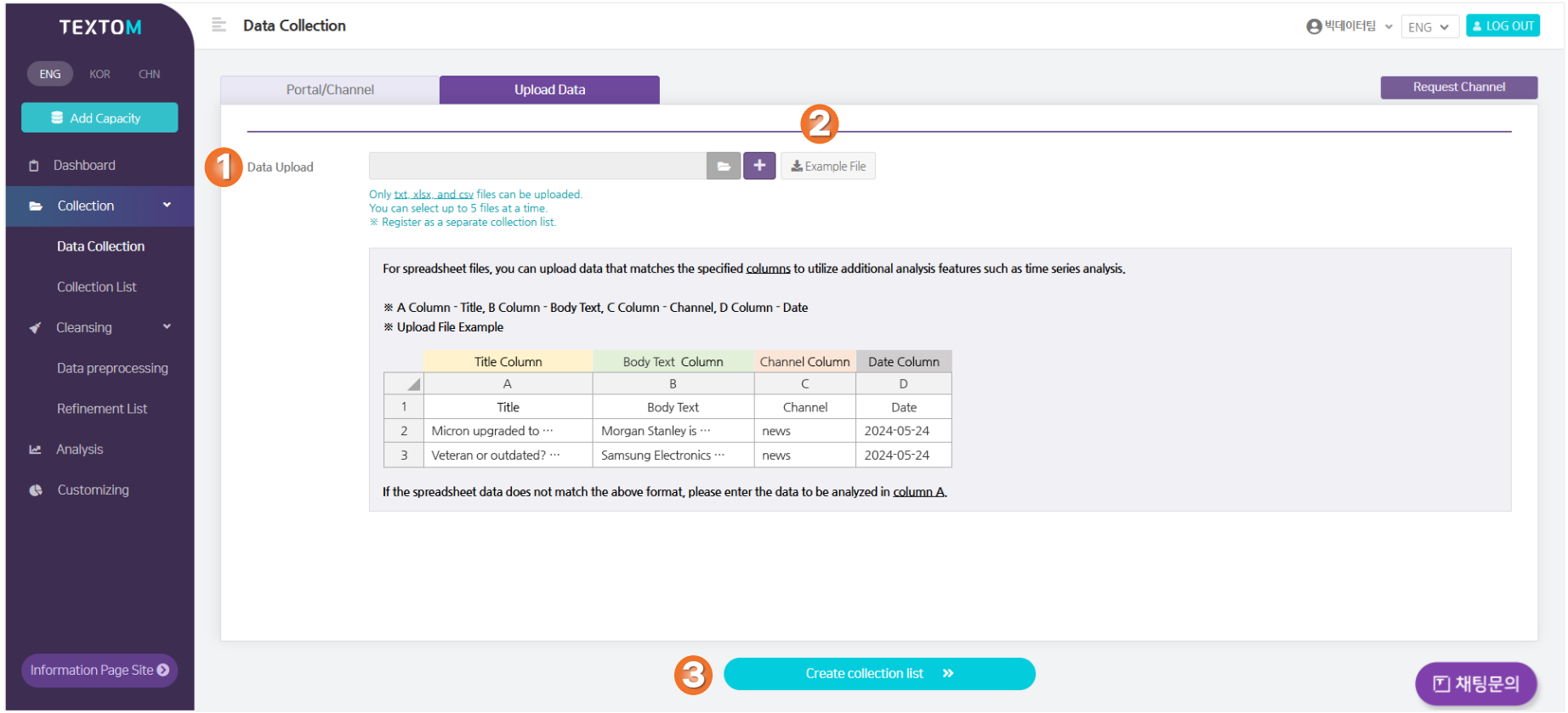


① Collect Keywords
Enter the keywords you want to collect. Use the [+], [-] buttons to create keyword sets and **search operators** to improve data quality.
Tip. Instead of broad keywords like "economy" and "tourism," use specific ones like "California economy" and "Napa Valley tourism" for better data quality and reduced usage.

② Collection Period
Set the desired data collection period. Yahoo, Quora, and Reddit do not support setting a collection period.
Tip. Set the collection period for up to 3 years. For longer or larger collections, please request through 'Request Channel'.

③ Channel
Select the collection portal/channel.
Tip. If you select Google channel, you cannot collect data alongside other channels. For Yahoo, Quora, and Reddit, please set up separate collection requests.

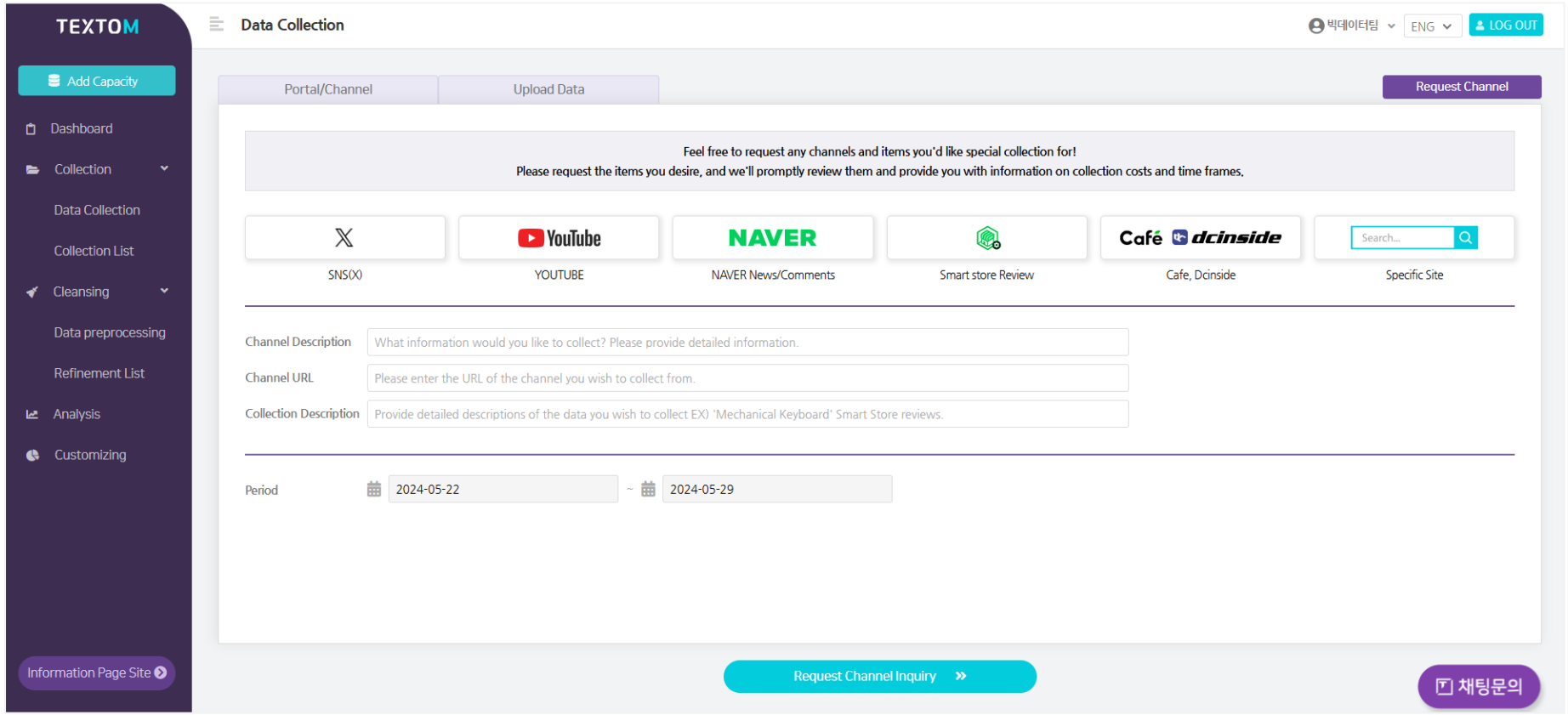
④ Run the data collection according to the selected options.



① **Data Upload**
Upload your own data. Click the [+] button to select up to 5 files.
Tip. To use additional analysis features like time series analysis, upload Excel files in the sample file format. Otherwise, enter the data for analysis in Column A.

② **Sample Format**
Click the button to download the sample format. Please refer to it and upload your data in the specified format.

③ Run the data collection according to the selected options.



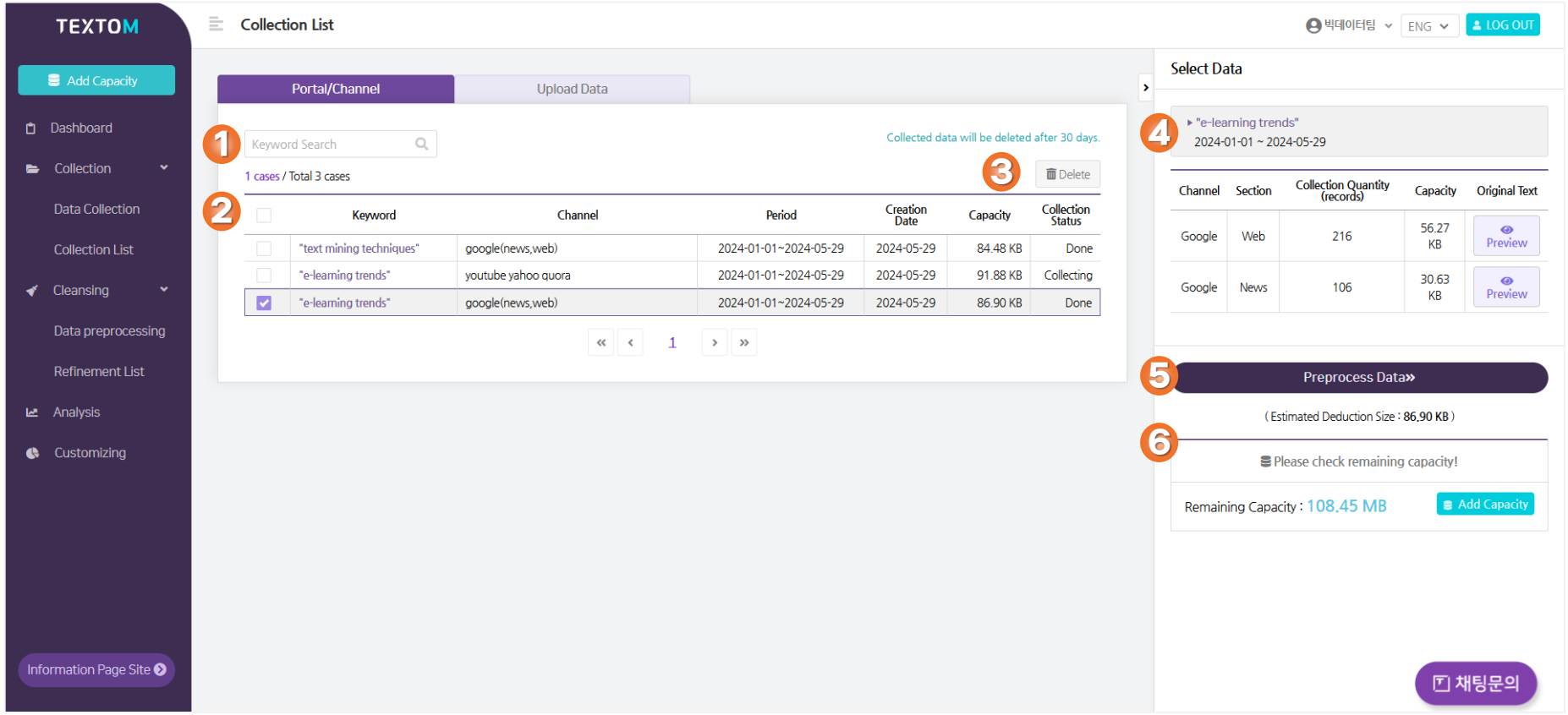
Data Collection Request

If you want to collect data outside of the channels provided by TEXTOM, please register as a request channel.

- Based on the contents of the request, we will check whether the data can be collected from the requested site and provide a quotation according to the calculation method.
- The collection method through the request channel is to receive only the collected data or to build it within the request channel page.
- The collected data is delivered as an EXCEL(.xlsx) file or txt file.
- Once built, you can use it in the same way as the existing channels provided in the collection service.

Collection List

Collected data from portals and channels.



① You can search using keywords.

② View Collected data
Shows the list of collected data. Click the checkbox to view the information of the corresponding data in the bottom section.

③ Delete Data
Select the checkbox and click the [Delete] button to remove the data from the collection list.
Tip. Data will be automatically deleted 30 days after collection. It is recommended to download the data that the user wants to keep.

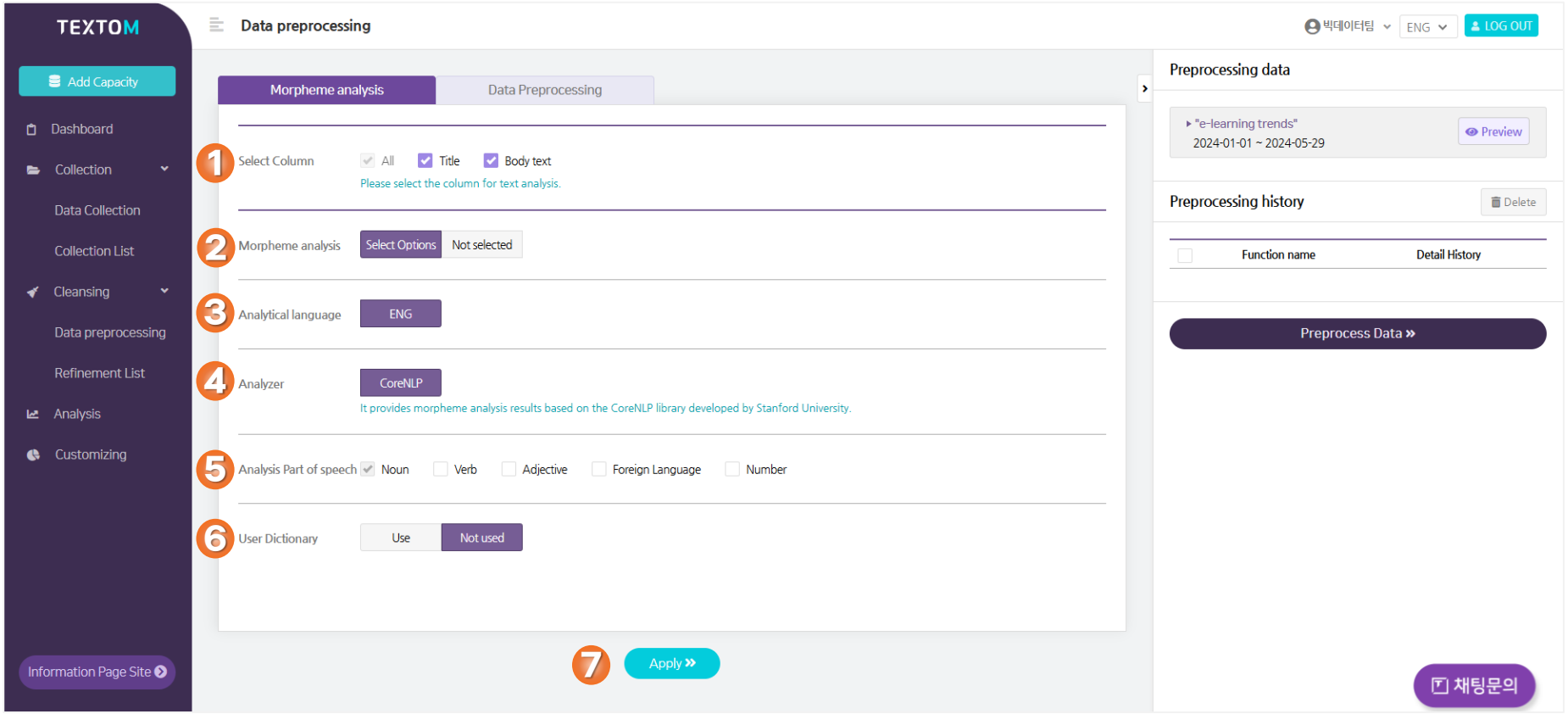
④ Check Data Information
Clicking on the data will show the channel, section, and original text in the preview area.
Tip. Previewing the original text helps in filtering spam keywords during the data processing stage.

⑤ Process the data.

⑥ You can check the estimated user usage and residual usage and add usage.

Cleansing Data

Select the morphological analysis options.



① **Select Analysis Column**
Select the column to analyze.
Tip. Analyze "title" for quick trends; analyze entire text for in-depth insights.

② **Morphological Analysis**
If you are uploading already processed data or analyzing the raw text as it is, select [Not selected].

③ Select a language for analysis.

④ TEXTOM provides morphological analysis results based on the CoreNLP library created by Stanford University.

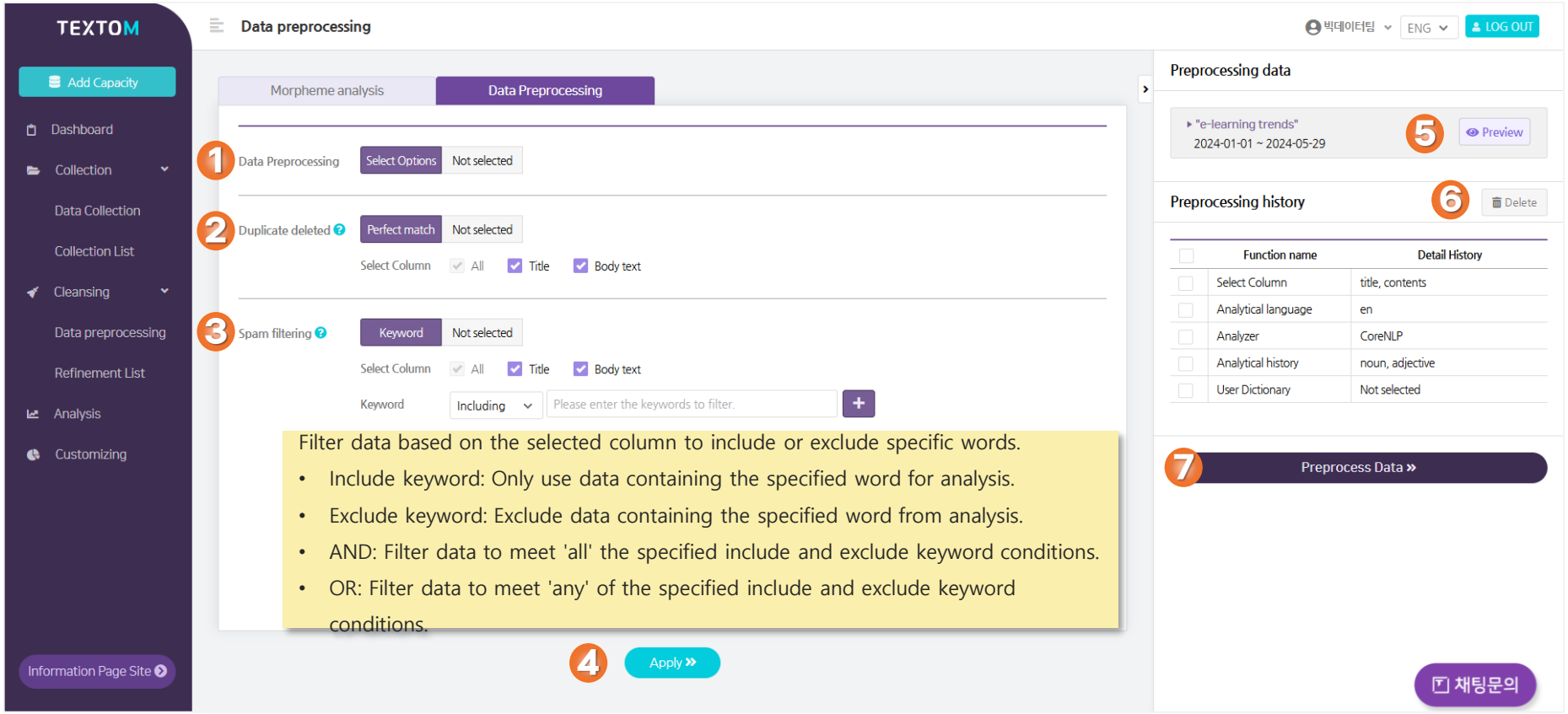
⑤ Select Part-of-Speech(POS) tagging criteria.
Tip. Verbs : interaction between subjects Adjectives : sentiment analysis Numbers : proper noun differentiation

⑥ **User Dictionary**
Use the user dictionary for efficient repeated analysis of specific topics or keywords.
Tip. Register keywords to refine in the user dictionary first. (My Page -> User Dictionary)

⑦ Click the [Apply] button to save the selected options in the preprocessing history.

Cleansing Data

Select the date processing options.



① **Data Cleansing Option Selection**
You can use cleansing features such as duplicate removal and spam filtering.
Tip. It is recommended to use these features to reduce data volume and enhance analysis efficiency.

② **Duplicate Removal**
Remove duplicate data based on the selected column(s). When multiple conditions are set, identical data within each column will be removed.

③ **Spam Filtering**
Filter data using keywords to include or exclude records based on the selected column.

④ Click the [Apply] button to save the selected options in the preprocessing history.

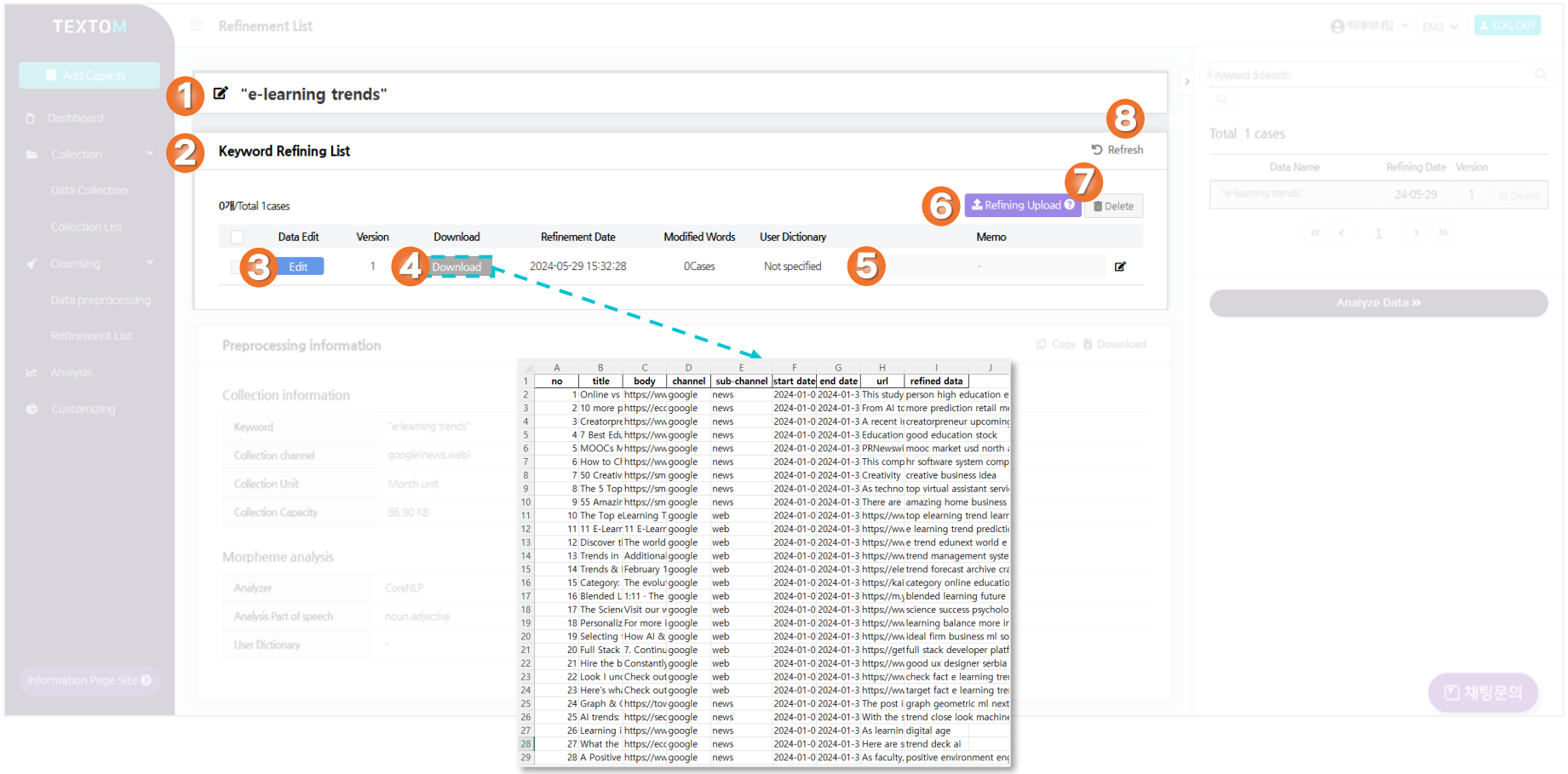
⑤ **Data Preview**
Preview the data to be analyzed.
Tip. Reviewing the original text beforehand aids in spam keyword filtering and helps reduce data usage.

⑥ Check the box and click [Delete] to remove the analysis options.

⑦ Proceed with data preprocessing with the selected analysis option.

Cleansing Data

Check and manage the keyword refinement status.



① Click the [Edit] button to modify the data name.

② Keyword Refinement List
You can check the history of keyword editing.

③ Real-Time Edit
Correct keywords directly on the web in real time without uploading a file.

④ Data Download
Download an Excel (.xlsx) file with raw data, refined data, dates, and related information.

⑤ Leave a brief note on the refinement version and history.

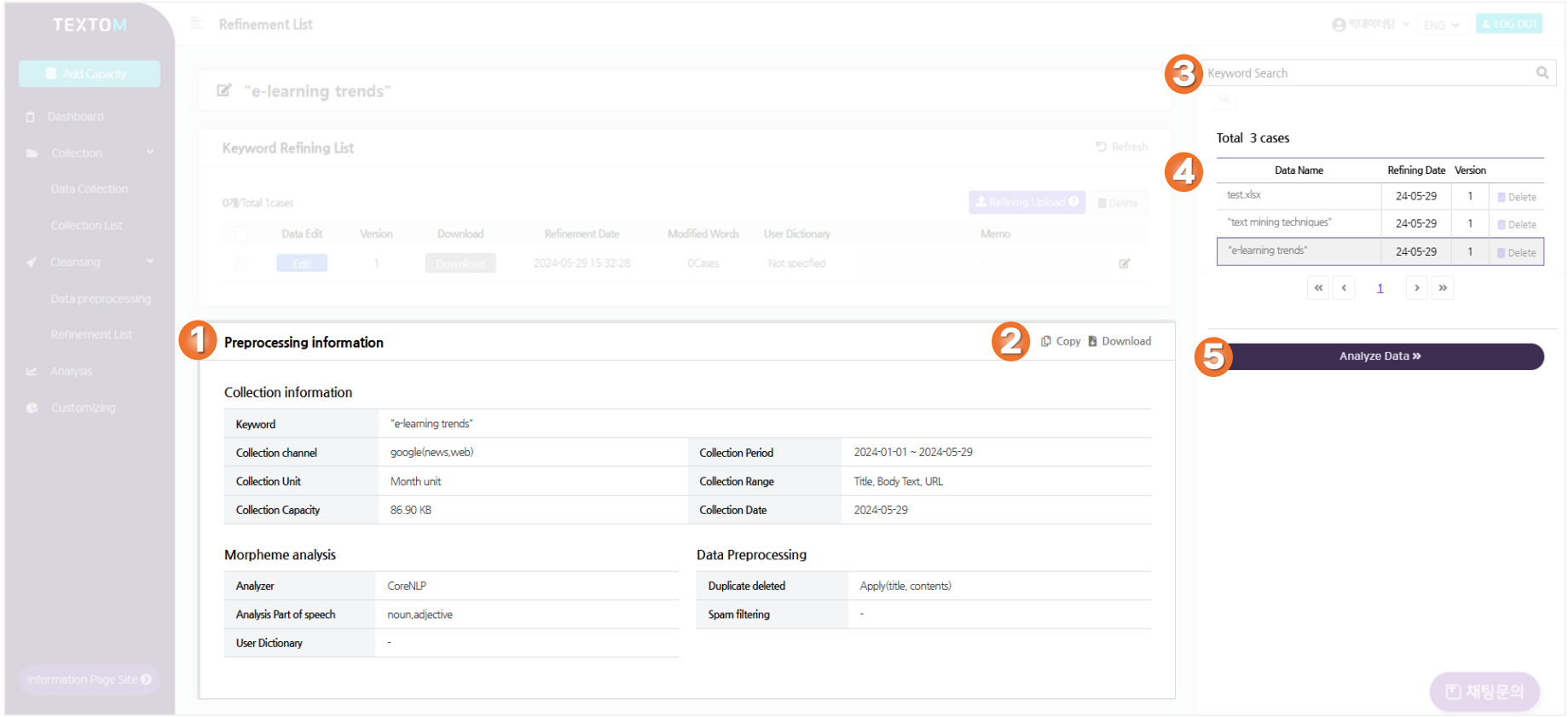
⑥ Upload Refined File
You can upload a new version by downloading the file in ④ and refining it yourself.
Tip: Ensure the file format matches the one downloaded in ④.

⑦ Delete Refinement History
Select the refinement history to delete using the checkbox, then click [Delete].

⑧ Refresh
Checks the refinement status in real time.

Cleansing Data

Manage the refinement list and data.



① **Processing Information**
View information on data collection and refinement.

② Copy or download the preprocessing information in table format.

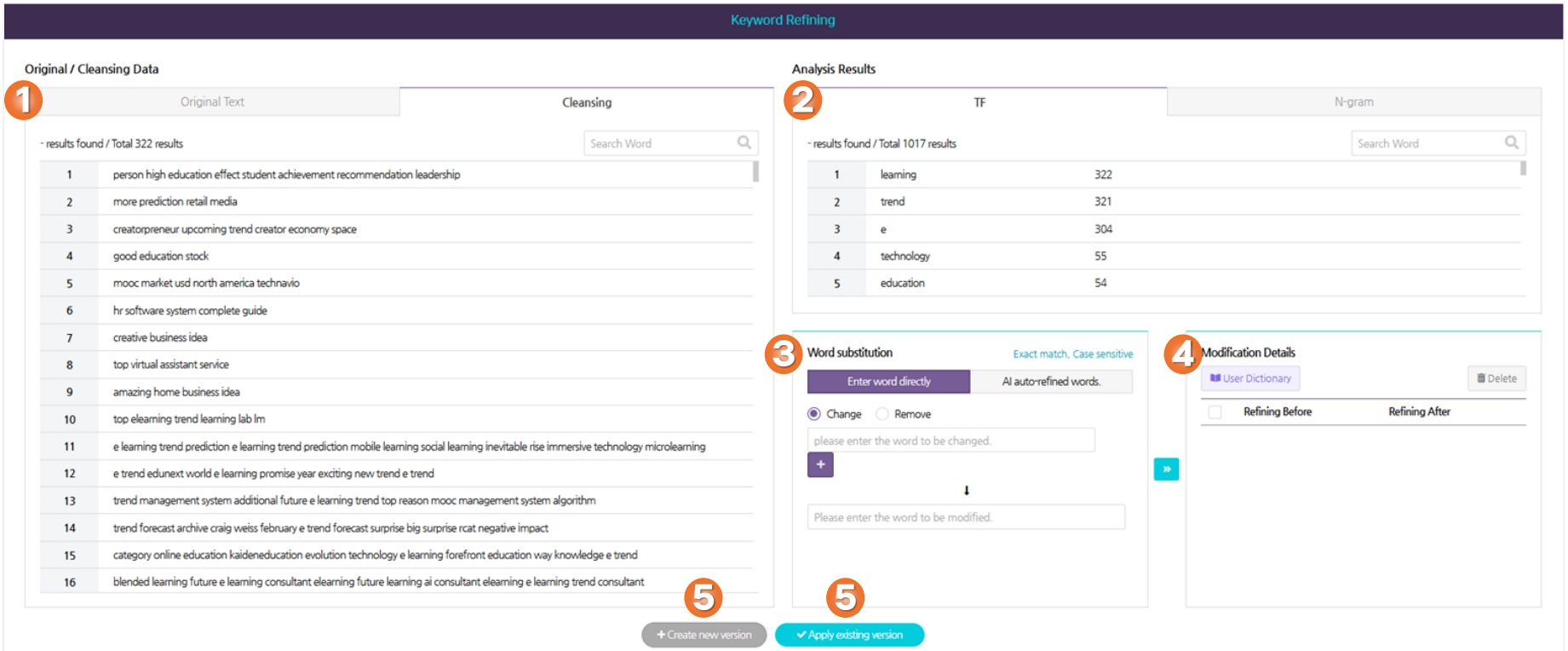
③ **Keyword Search**
Search the refinement list by keyword.

④ **Refinement List Management**
Perform management tasks such as viewing and deleting the refinement list.

⑤ Analysis is performed on the selected data.

Cleansing Data

Refine keywords directly on web pages, without needing to upload data.



① **Raw/Refined Data Tab**
Switch between the raw data tab and the refined data tab to view the data.

② **View Analysis Results**
View TF (Term Frequency) and N-gram analysis results for keyword refinement.

③ **Edit Keywords**
Input and change keywords directly or use the [AI automatic refinement words] function.

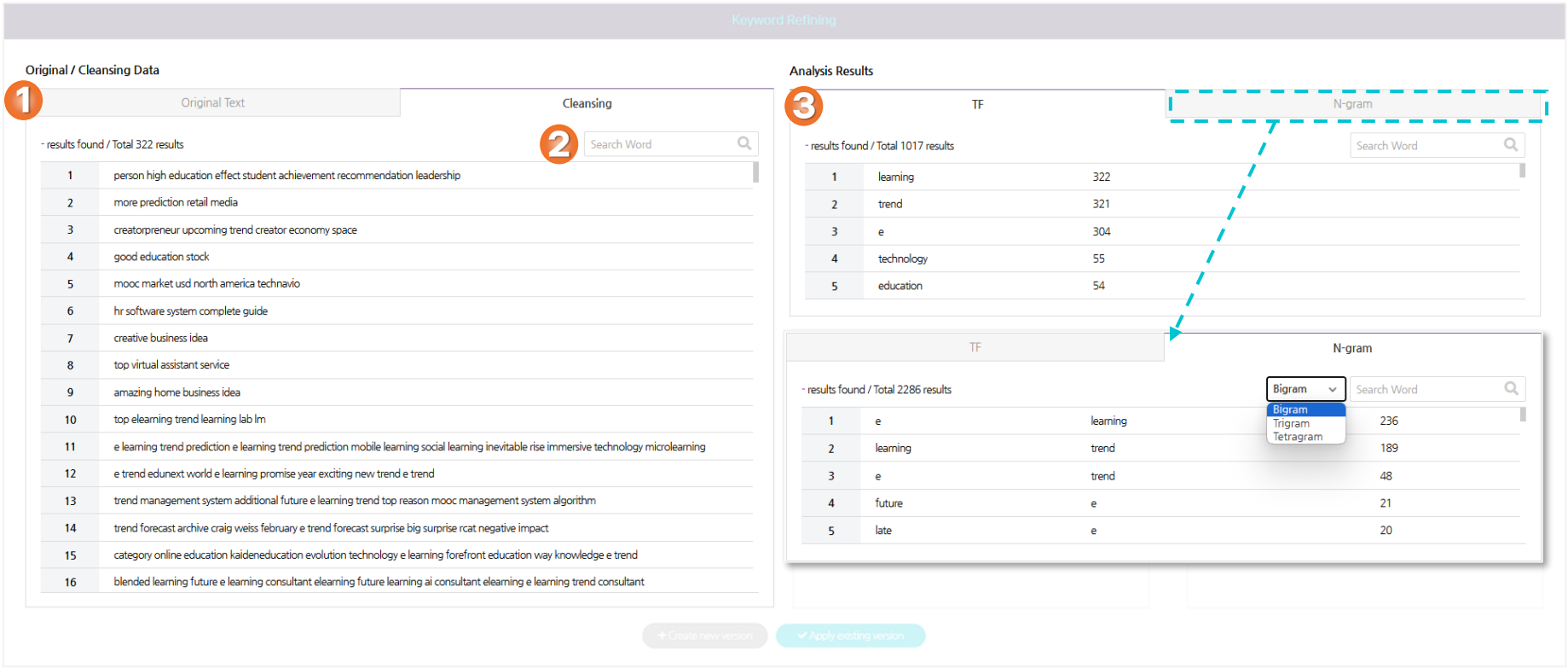
④ **Edit History**
View and delete keyword edit history. Register the current edit history in the user dictionary or load an existing user dictionary.

⑤ **Create New Version**
Add the edit history to a new version.

⑥ **Apply Previous Version**
Apply the edit history to a previous refinement version.
Tip. Once applied, changes cannot be undone. Review edit history and download refined data in advance.

Cleansing Data

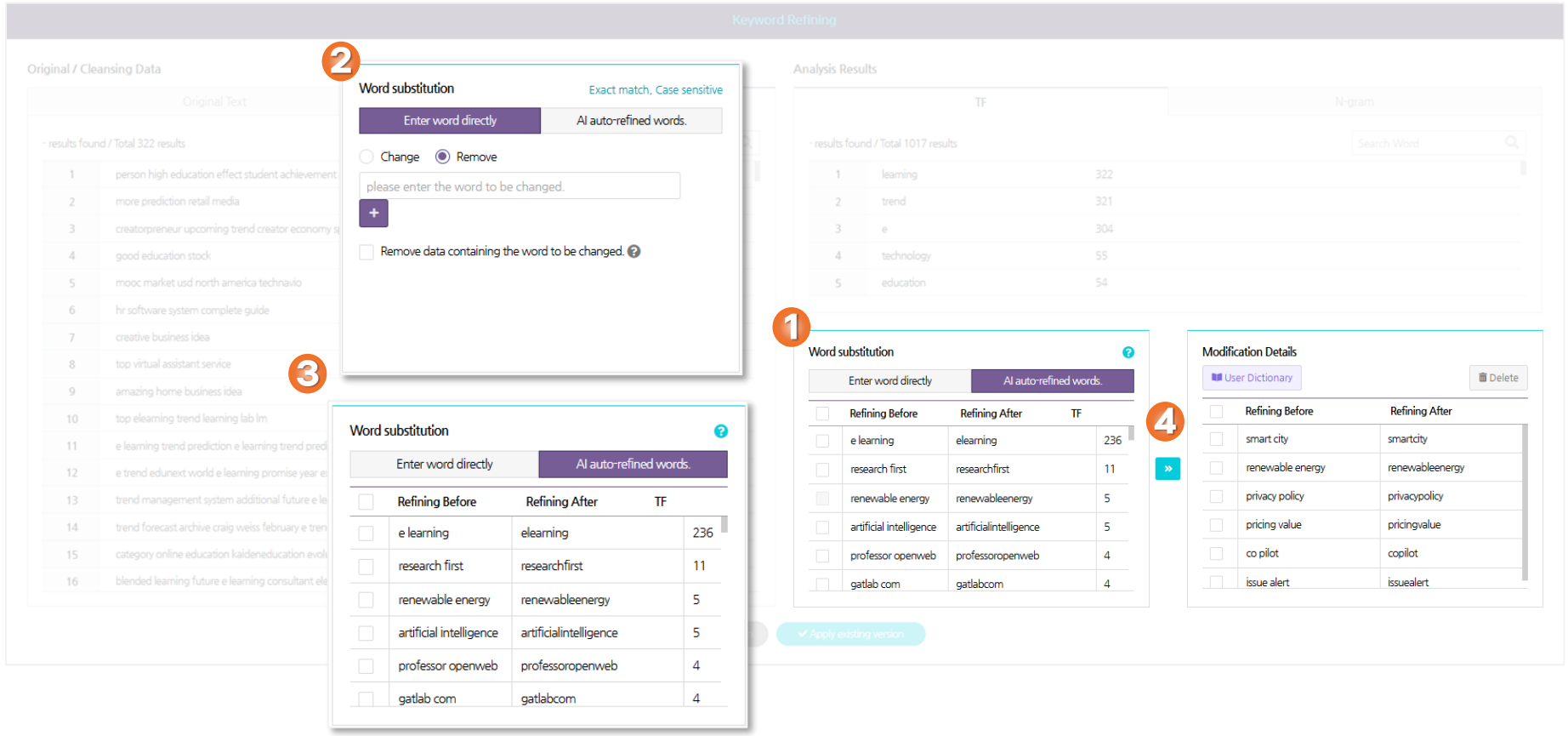
Refine your keywords based on the raw/refined data and the results of your analysis.



① **Raw/Refined Data Tab**
Switch between the raw data tab and the refined data tab to view the data.

② **Keyword Search**
Analyze how specific keywords are used in the raw and refined data, and determine which words are frequently used together with the keywords through N-gram analysis. This can be useful for keyword refinement.
Tip. Use this functionality during the data refinement process to verify data connections, linked words, and the outcomes of applied refined keywords.

③ **View Analysis Result**
View TF (Term Frequency) and N-gram analysis results.
Tip. TF and N-gram analysis results can help adjust, remove, or identify relevant keywords effectively.



① **Direct Word Refinement**
Directly input the word to be changed and the replacement word. If replacing multiple words with one, click the [+] button to add words.

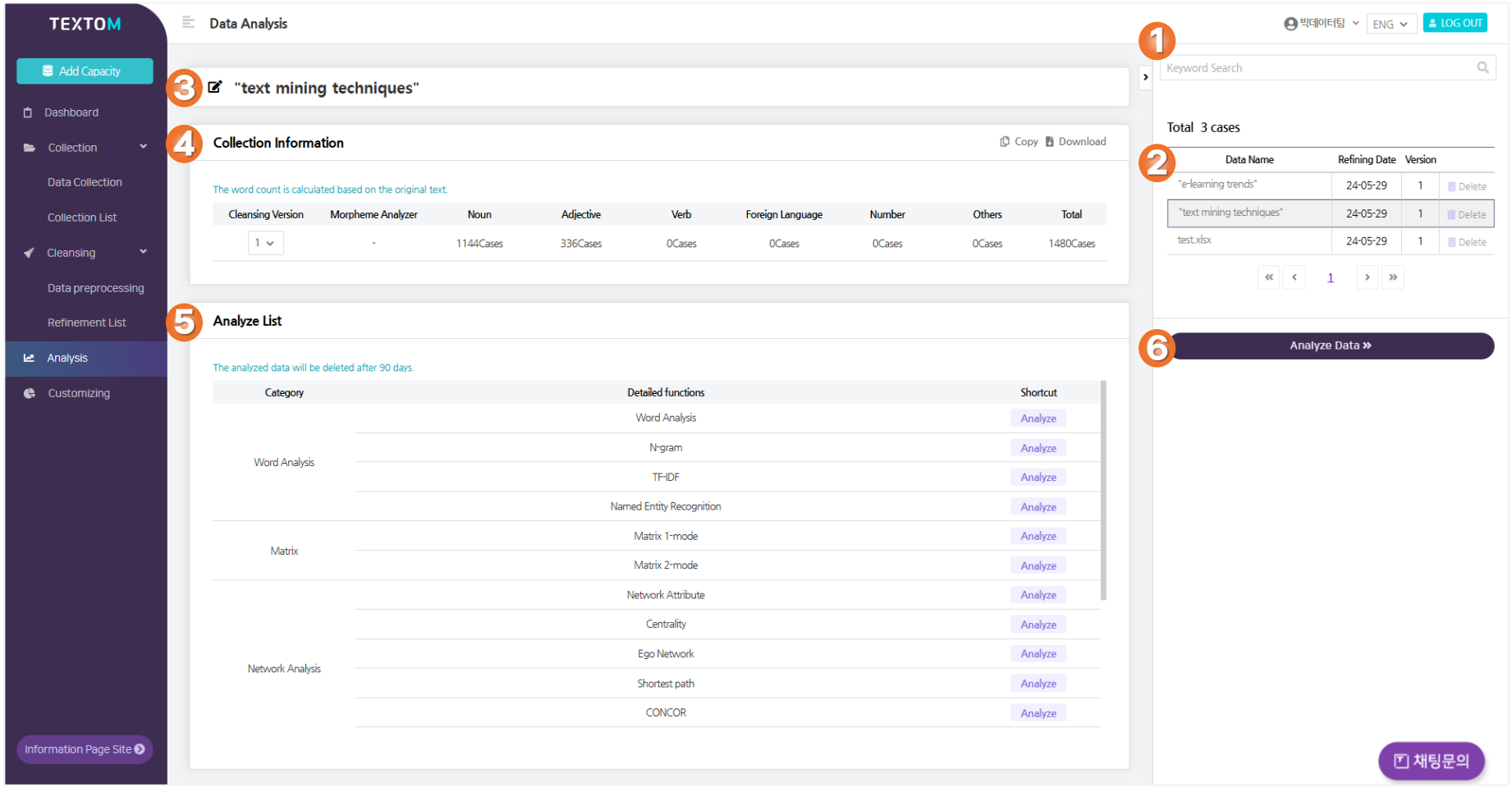
② **Keyword and Data Removal**
Remove specific words from the entire dataset. Add words to remove by clicking the [+] button. To remove data containing specific keywords, select the checkbox.

③ **AI Automated Refinement**
Calculate the probability of words appearing consecutively based on AI, recommending up to 200word pairs.

④ Apply the changes to the modification history by pressing the button.

Data Analysis

View the analysis list and check the details of the data.



① Search for keywords to view data containing those words.

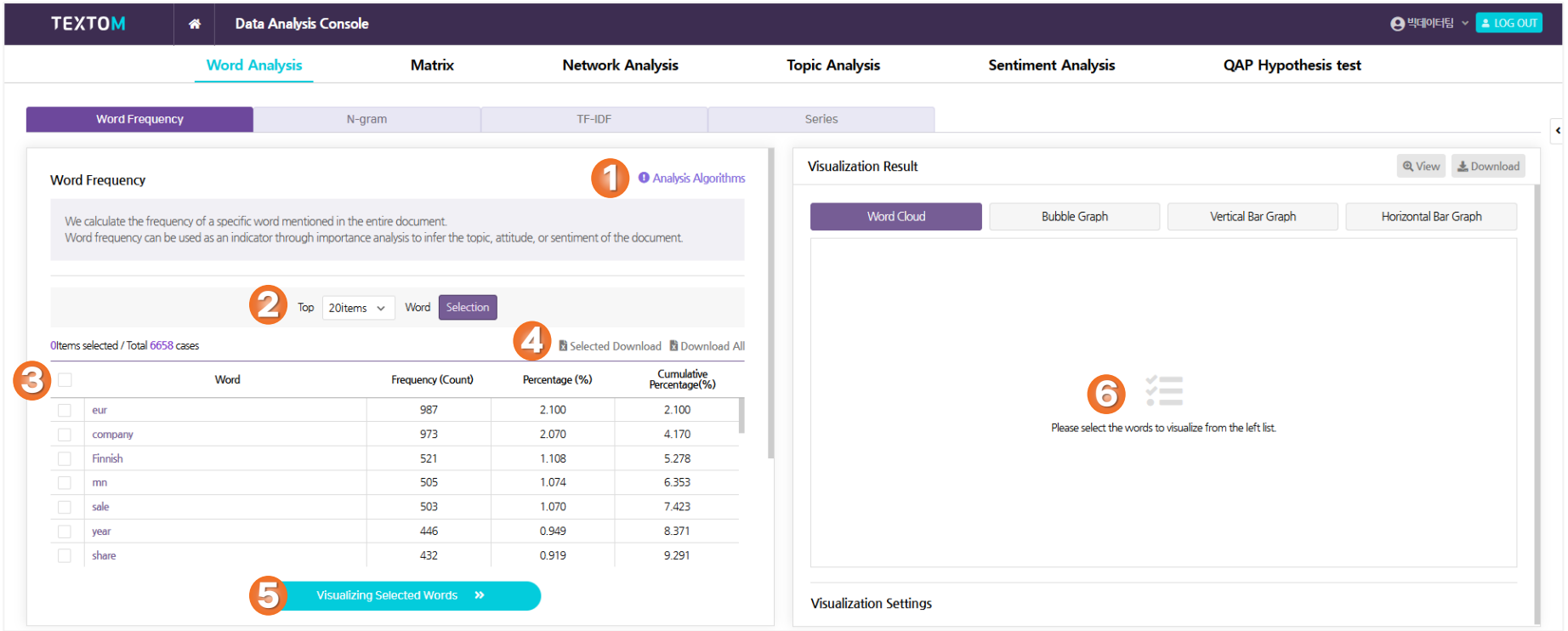
② **Analysis List**
View and delete data that has been refined. Clicking on specific data displays its information in the left area.
Tip. Click the [>] button to hide the analysis list window.

③ **Modify Data Name**
Click the [Edit] button to easily change the data name to your preferred title.

④ **Analysis Information**
View the morphological analysis results of the data.
Tip. You can select the refined version and copy the analysis information or download it as an Excel (.xlsx) file.

⑤ **Analysis History**
Access the analysis results by clicking the "Analyze" button.

⑥ Navigate to the analysis results screen for the selected data.



① **Analysis Algorithm**
Click the button to view the description of the analysis algorithm.

② **Top Words Auto-Select**
Automatically select the top 20, 50, 75, or 100 words from all words.

③ **Word Selection**
Click the top checkbox to automatically select the top 100 words. To manually select words, use the checkbox next to each word.
Tip. The percentage represents the amount mentioned when the total mentions are scaled to 100.

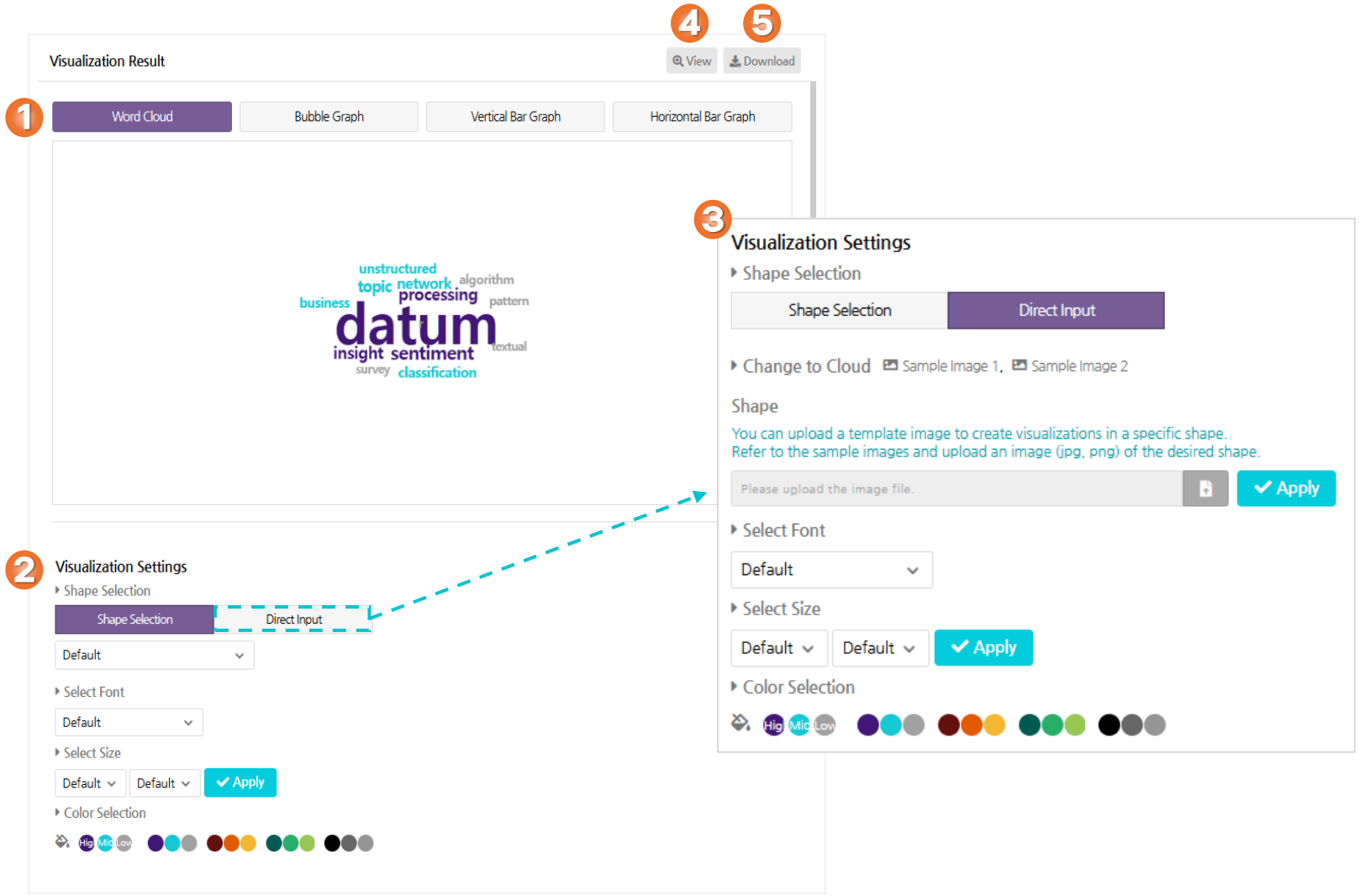
④ **Download Analysis Results**
Click [Selected Download] or [Download All] to download word frequencies and percentages into an Excel (.xlsx) file.

⑤ Visualize the selected words.
Tip. For better visibility, it is recommended to select between 50 to 70 words.

⑥ Displays the visualization results.

Data Analysis

Displays the visualization results of Word Frequency (TF).



① **Visualization Results**
Provides four types of visualization results.

Word Cloud	Bubble Chart
Vertical Bar Chart	Horizontal Bar Chart

② **Visualization Setting**
Users can modify the attributes of the visualization results as desired.

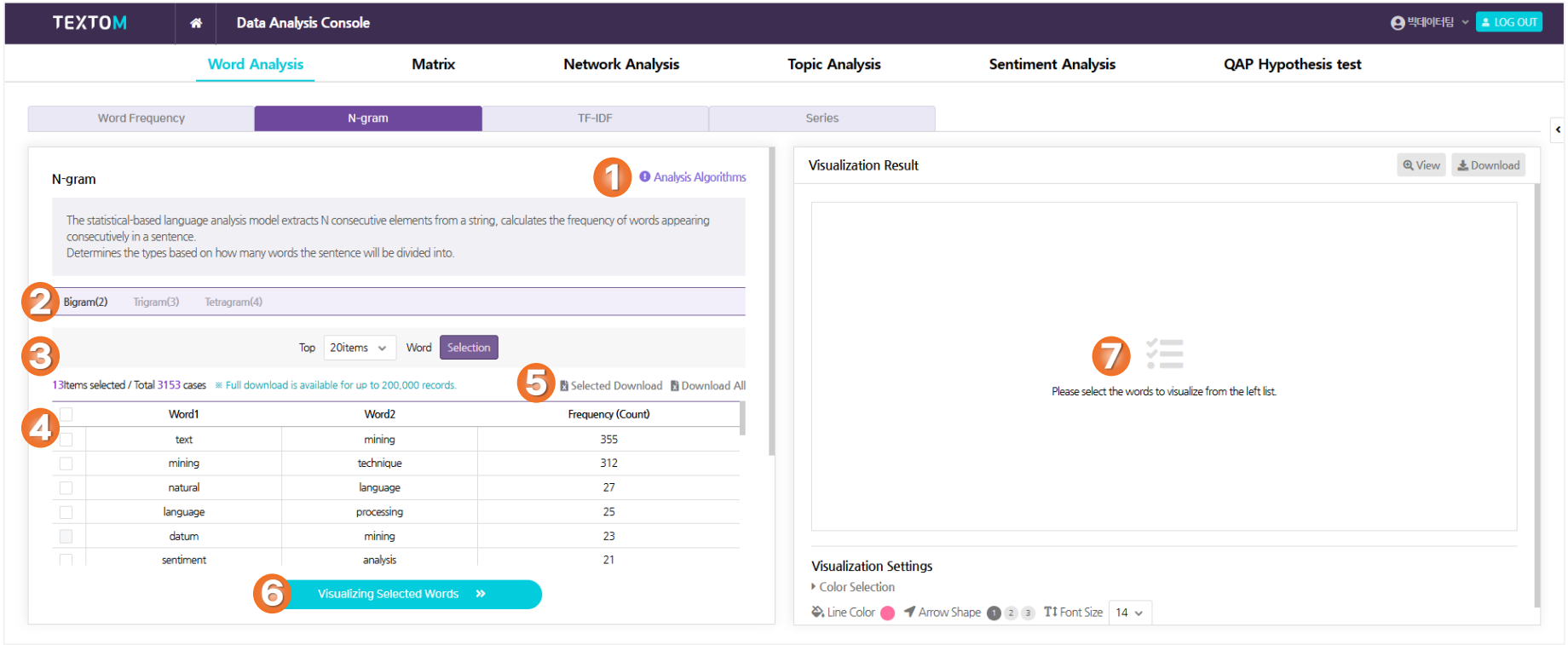
③ **Shape Selection (Word Cloud)**
Enables users to select a specific shape for the word cloud visualization.
Tip. Refer to sample images and upload an image of the desired shape.

④ **Zoom In**
Click [View] to display an enlarged version of the selected visualization in a pop-up window.

⑤ **Download Visualization**
Download the visualization results as a PNG file.

Data Analysis

The results of an N-gram analysis.



① **Analysis Algorithm**
Click the button to view the description of the analysis algorithm.

② **Set N-gram Range**
Select the number of consecutive words from 2 (Bigram) to 4 (Tetragram).

③ **Top Words Auto-Select**
Automatically select the top 20, 50, 75, or 100 words from all words.

④ **N-gram Word Selection**
Click the top checkbox to automatically select the top 100 words. To manually select words, use the checkbox next to each word.

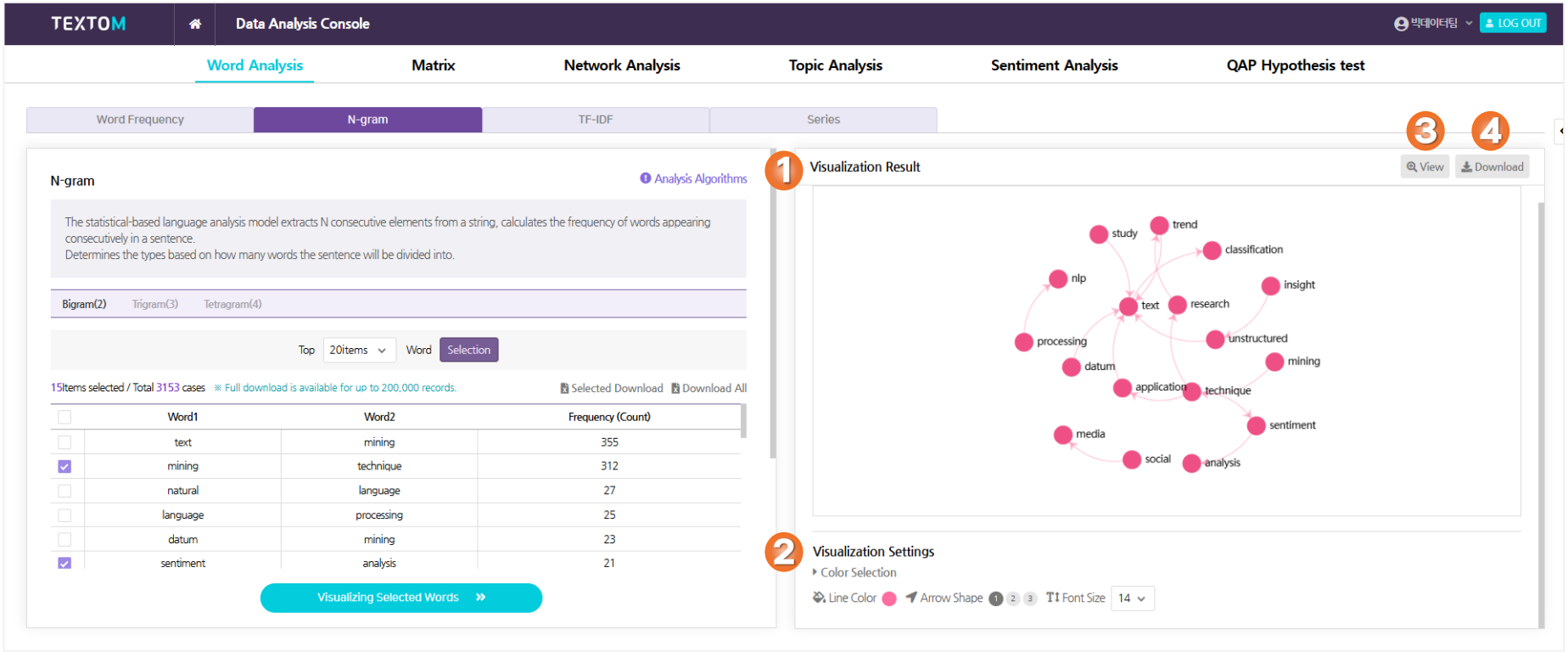
⑤ **Download Analysis Results**
Click [Selected Download] or [Download All] to download word frequencies and percentages into an Excel (.xlsx) file.

⑥ Visualize the selected words.

⑦ Displays the visualization results

Data Analysis

Displays the visualization results of N-gram.

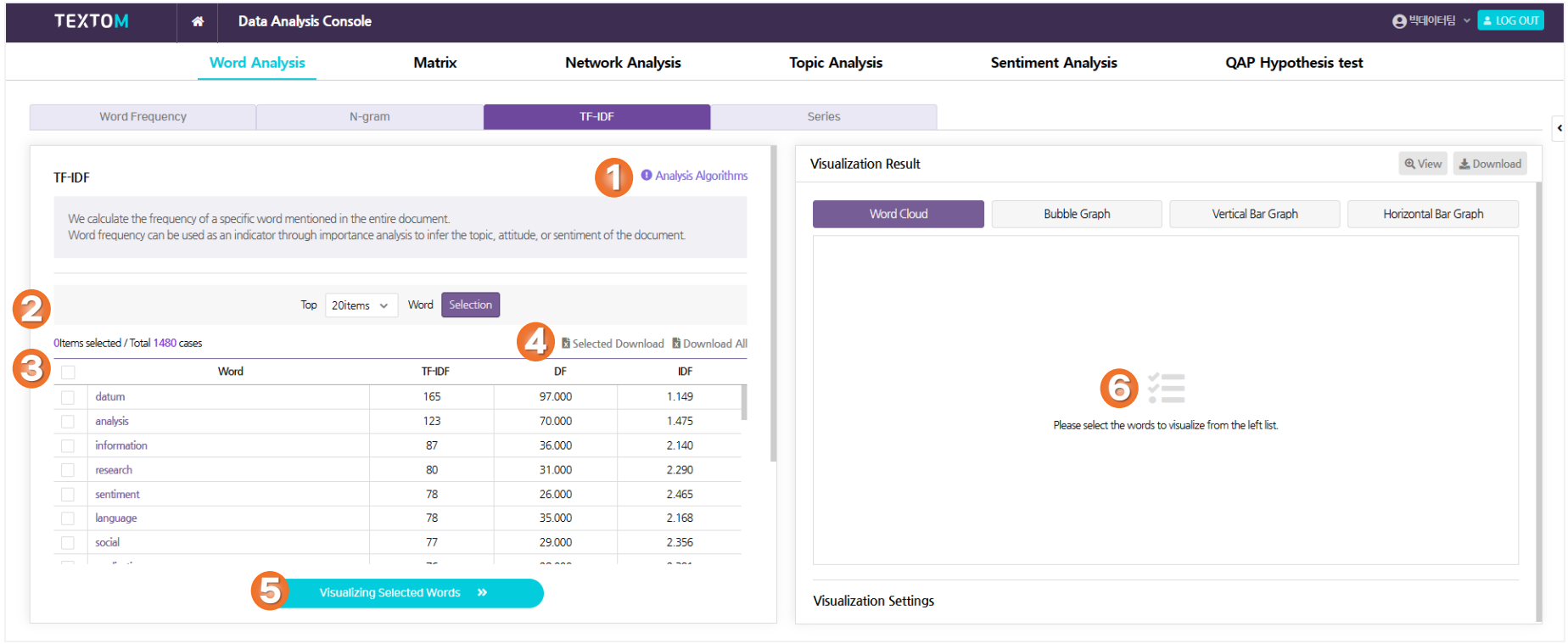


① **Visualization Result**
The results are provided as a directional network graph.
Tip. It is important to select relevant words according to the analysis purpose, and it is advisable to exclude common words that do not contain relevant information.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

② **Visualization Setting**
Users can modify the attributes of the visualization results as desired.

③ **Zoom In**
Click [View] to display an enlarged version of the selected visualization in a pop-up window.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

④ **Download Visualization**
Download the visualization result as a PNG file.



① **Selected Algorithm Description**
Click the button to view the description of the selected algorithm to analyze.

② **Top Words Auto-Select**
Automatically select the top 20, 50, 75, or 100 words from all words.

③ **TF-IDF Word Selection**
Click the top checkbox to automatically select the top 100 words. To manually select words, use the checkbox next to each word.
Tip. The higher the TF-IDF value, the more important the word is within a specific document.

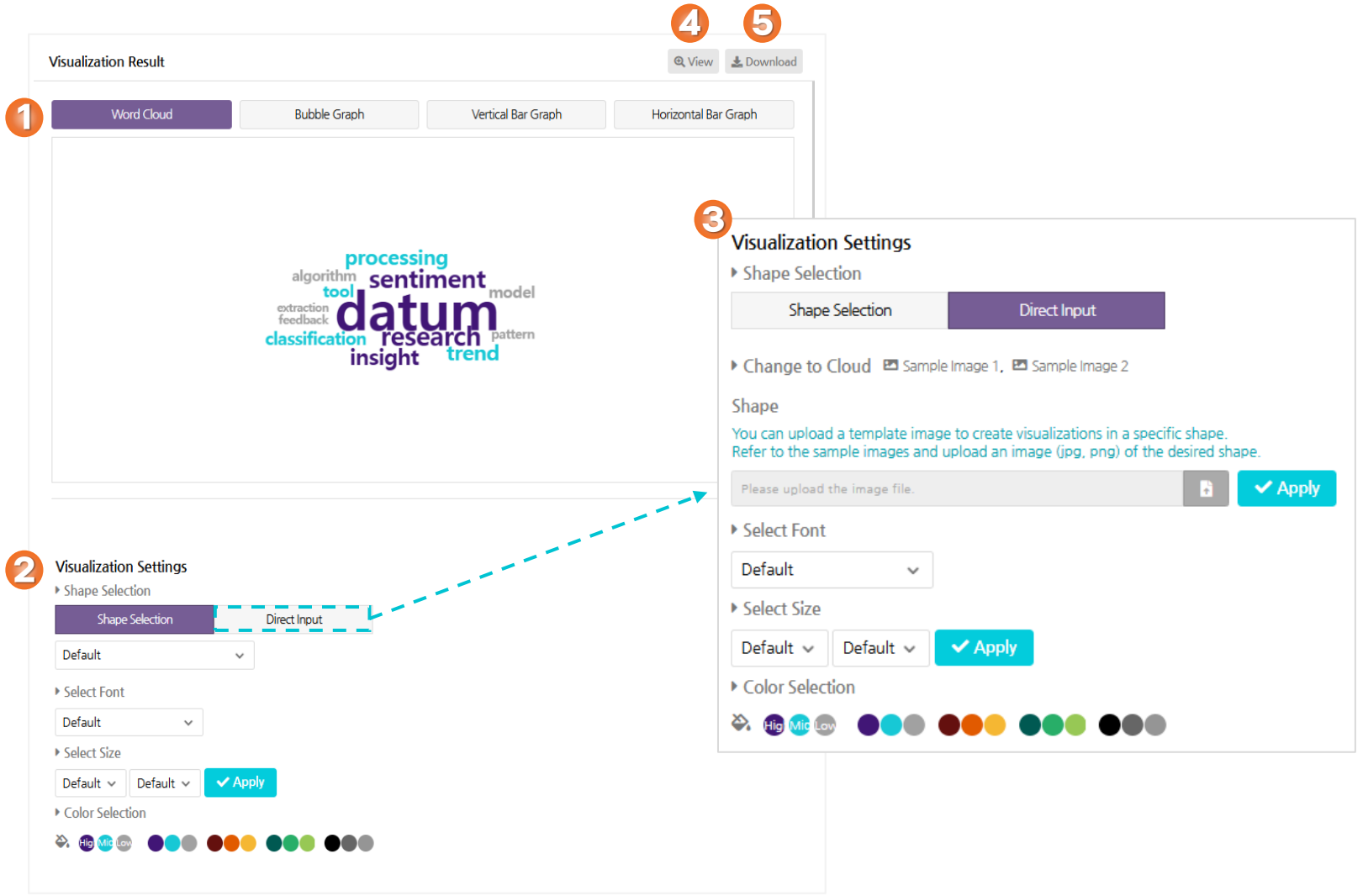
④ **Download Analysis Results**
Click [Selected Download] or [Download All] to download word frequencies and percentages into an Excel (.xlsx) file.

⑤ Visualize the selected words.

⑥ Displays the visualization results

Data Analysis

Displays the visualization results of TF-IDF.



① **Visualization Results**
Provides four types of visualization results.

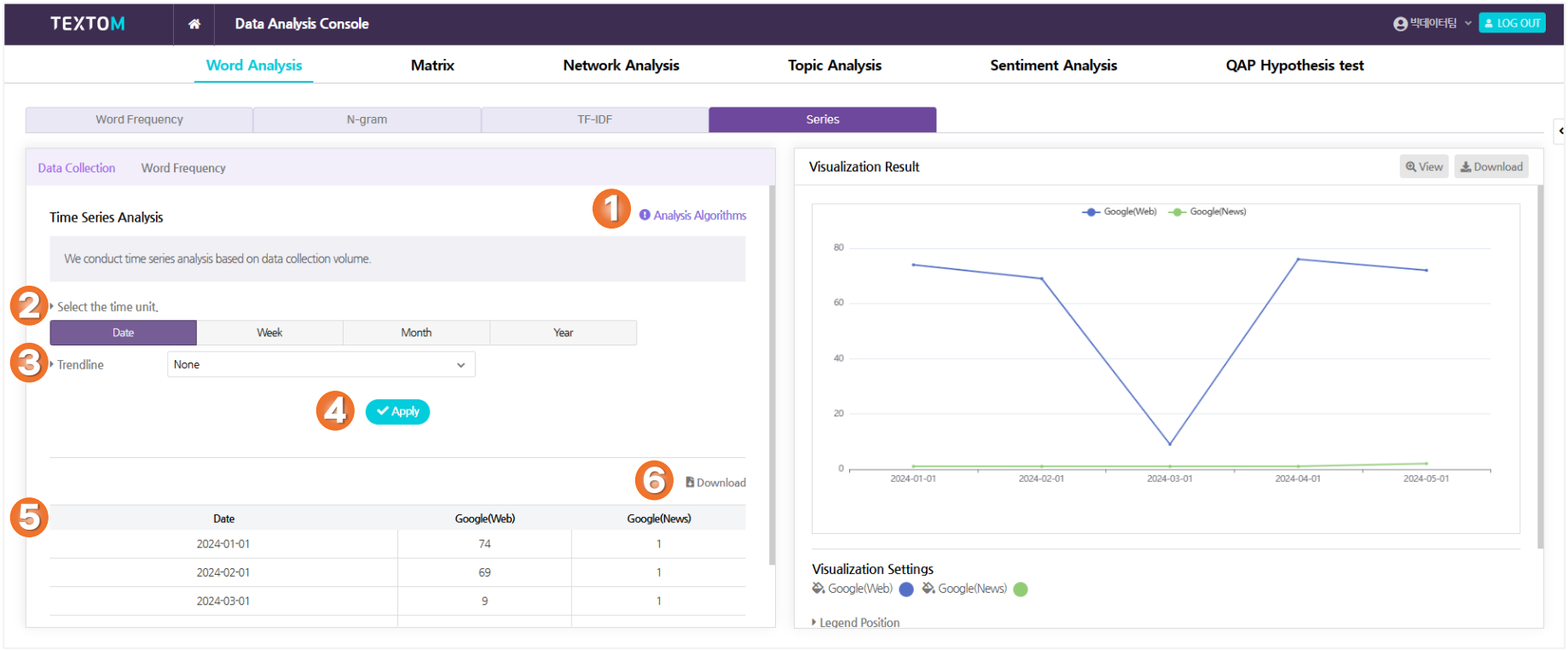
Word Cloud	Bubble Chart
Vertical Bar Chart	Horizontal Bar Chart

② **Visualization Setting**
Users can modify the attributes of the visualization results as desired.

③ **Shape Selection (Word Cloud)**
Enables users to select a specific shape for the word cloud visualization.
Tip. Refer to sample images and upload an image of the desired shape.

④ **Zoom In**
Click [View] to display an enlarged version of the selected visualization in a pop-up window.

⑤ **Download Visualization**
Download the visualization result as a PNG file.



① **Analysis Algorithm**
Click the button to view the description of the analysis algorithm.

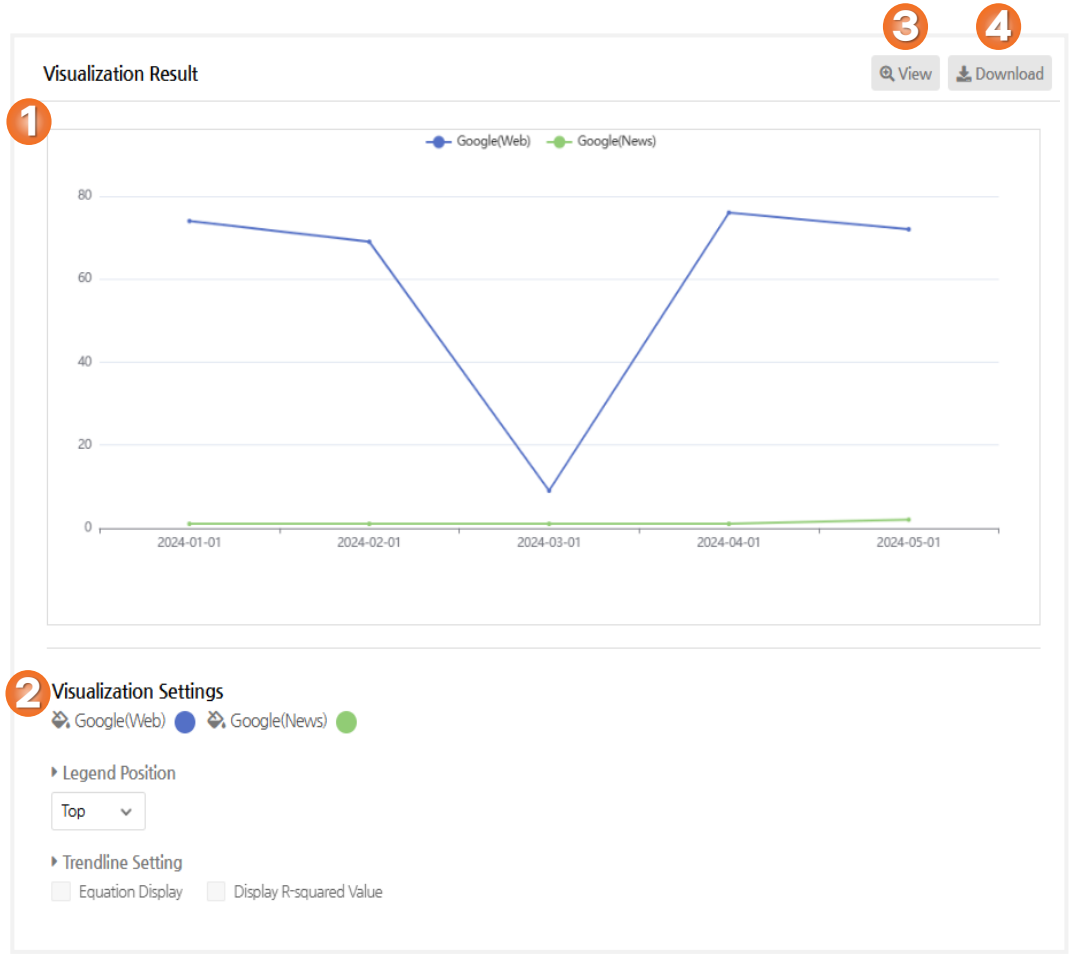
② **Select Time Unit**
Set the time unit to day, week, month, or year based on the data collection date.
Tip. This option is available when the 'Unit Collection' feature is used during the data collection stage.
Day : Day, Week, Month, Year
Week : Week, Month, Year
Month : Month, Year
Year : Year

③ **Display Trend Line**
Show a linear trend line in the visualization results to easily identify data collection trends.
Tip. To display the trend line in the visualization results, ensure you select [Linear].

④ Apply the selected options to the time series analysis.

⑤ Display the time series analysis results of the data collection volume.

⑥ Download the analysis result as an Excel (.xlsx) file.

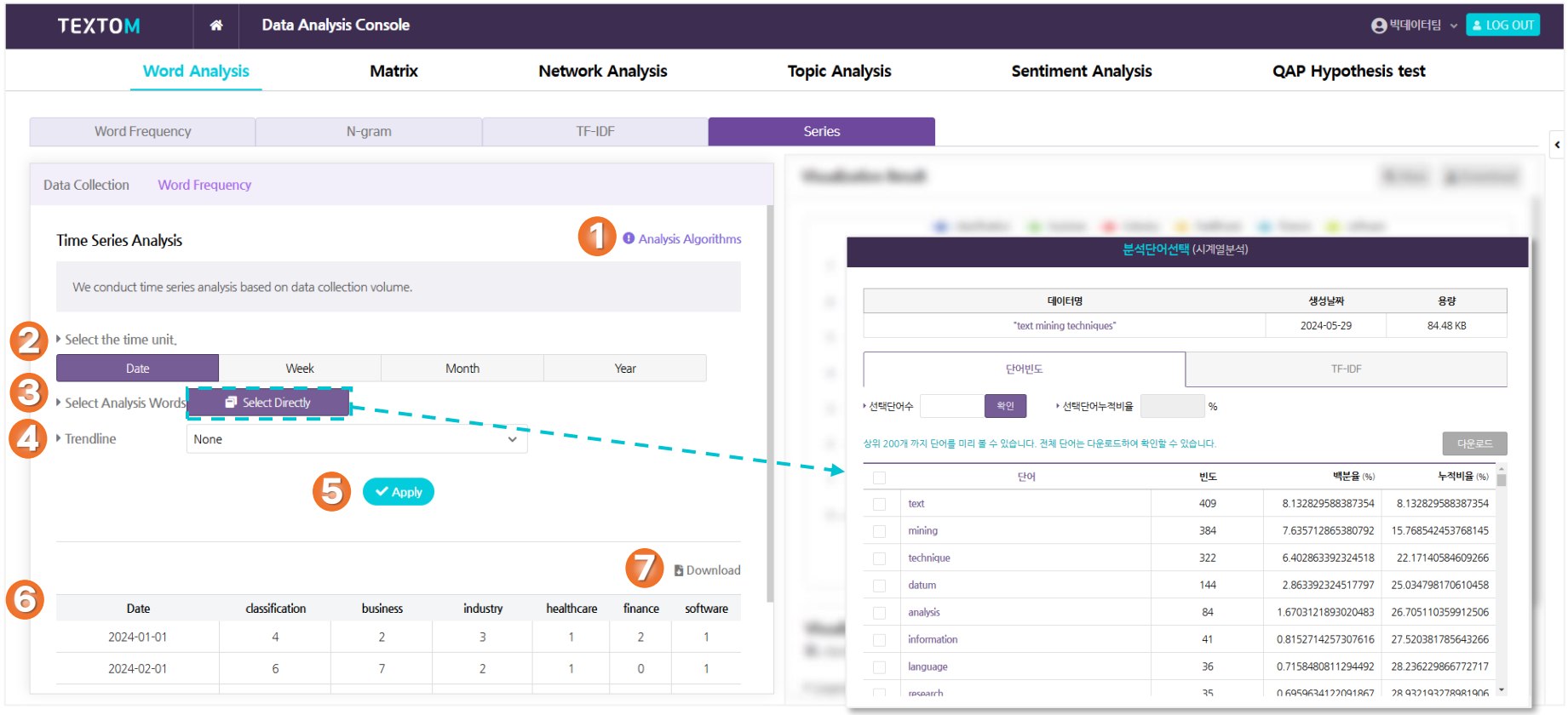


① Visualization Results
Displays the result of the analysis in a linear graph.

② Visualization Setting
Users can modify the attributes of the visualization results as desired.

③ Zoom In
Click [View] to display an enlarged version of the selected visualization in a pop-up window.

④ Download Visualization
Download the visualization result as a PNG file.



① **Selected Algorithm Description**
Click the button to view the description of the algorithm.

② **Select Time Unit**
Set the time unit to day, week, month, or year based on the data collection date.
Tip. This option is available when the 'Unit Collection' feature is used during the data collection stage.
Day : Day, Week, Month, Year
Week : Week, Month, Year
Month : Month, Year
Year : Year

③ **Select Word for Analysis**
select 2 to 10 words by clicking the "Select Now" button.

④ **Display Trend Line**
Show a linear trend line in the visualization results to easily identify data collection trends.
Tip. To display the trend line in the visualization results, ensure you select [Linear].

⑤ Apply the selected options to the time series analysis.

⑥ Display the time series analysis result of the data collection volume.

⑦ Download the analysis result as an Excel (.xlsx) file.

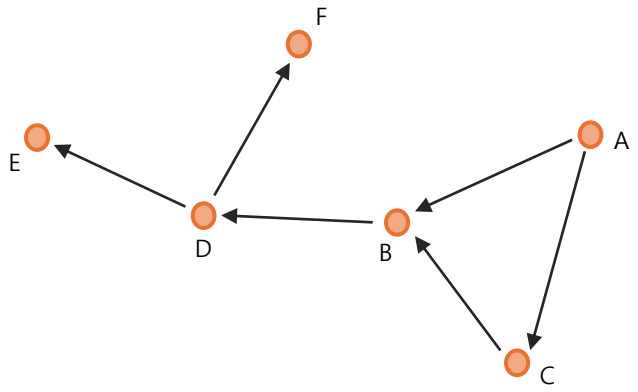
① Visualization Results
Displays the result of the analysis in a linear graph.

② Visualization Setting
Users can modify the attributes of the visualization results as desired.

③ Zoom In
Click [View] to display an enlarged version of the selected visualization in a pop-up window.

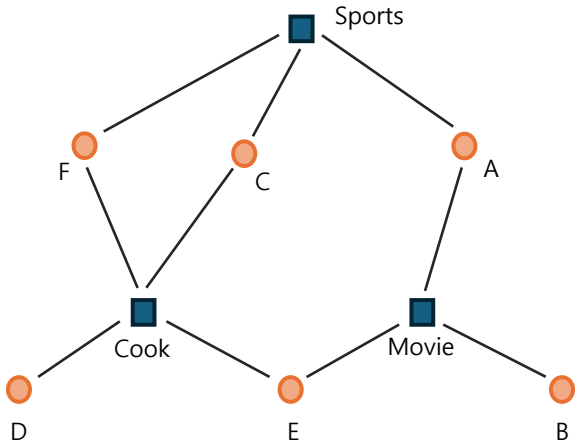
④ Download Visualization
Download the visualization results as a PNG file.

1 1-Mode Matrix



	A	B	C	D	E	F
A	0	1	1	0	0	0
B	0	0	0	1	0	0
C	0	1	0	0	0	0
D	0	0	0	0	1	1
E	0	0	0	0	0	0
F	0	0	0	0	0	0

1 2-Mode Matrix



	Sprots	Movie	Cook
A	1	1	0
B	0	1	0
C	1	0	1
D	0	0	1
E	0	1	1
F	1	0	1

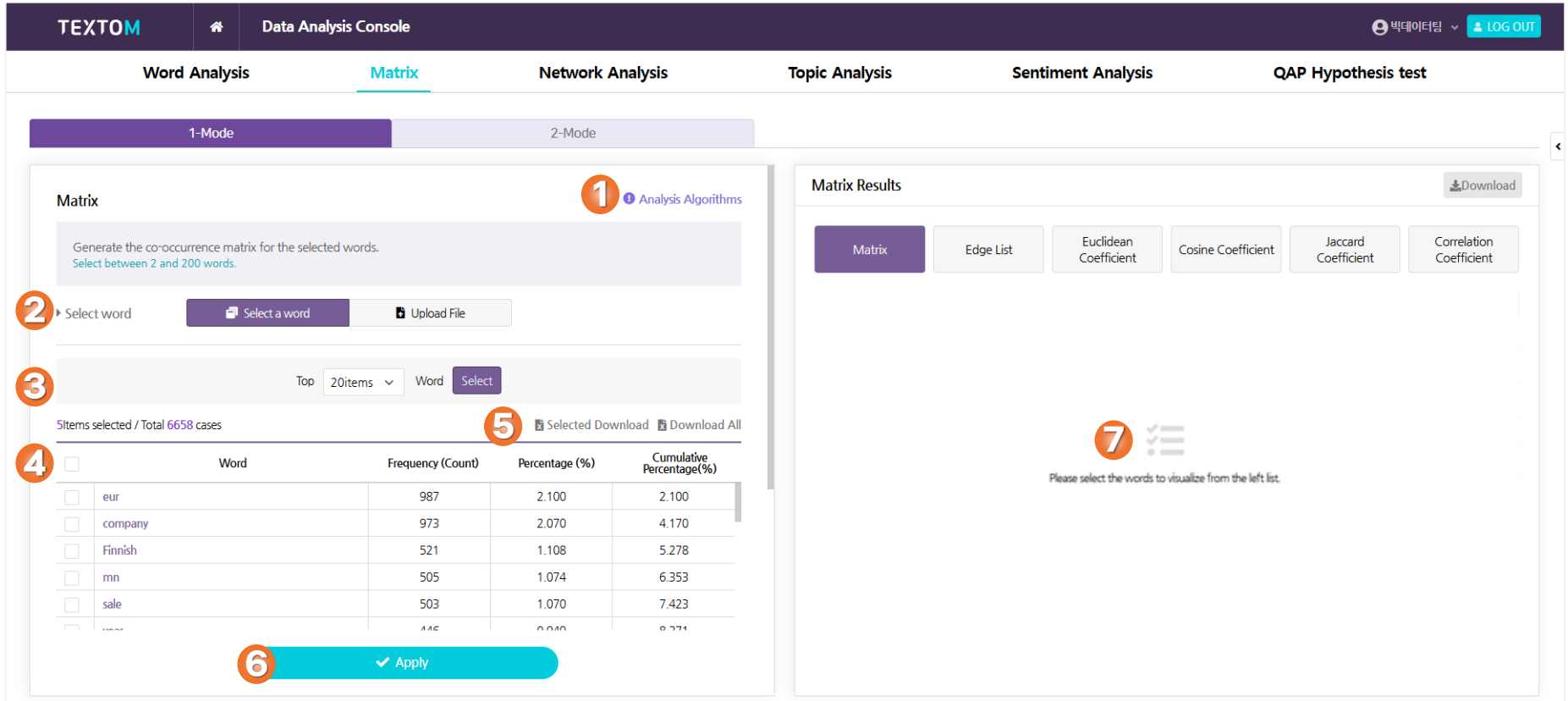
Matrix data is analyzed by constructing a matrix of word co-occurrence frequencies.

① 1-Mode Matrix
The most common type of network, with rows and columns representing the same words.
Tip. Suitable for analysis when the main node.

② 2-Mode Matrix
Rows and columns represent different words, allowing for analysis of relationships between different types of entities.
Tip. Suitable for analyzing relationships between different types of entities.

Matrix

Perform 1-Mode Matrix Analysis.



① **Selected Algorithm Description**
Click the button to view the description of the algorithm.

② **Set Word Selection Method**
Select 2 to 200 words for analysis. You can choose words directly from the analysis console or upload an Excel file.

③ **Top Words Auto-Select**
Automatically select the top 20, 50, 75, or 100 words from all words.

④ **Word Selection**
Click the top checkbox to automatically select the top 100 words. To manually select words, use the checkbox next to each word.
Tip. Selecting key words related to the analysis topic is crucial. Review the results after the initial selection and adjust as necessary to enhance the accuracy of the analysis.

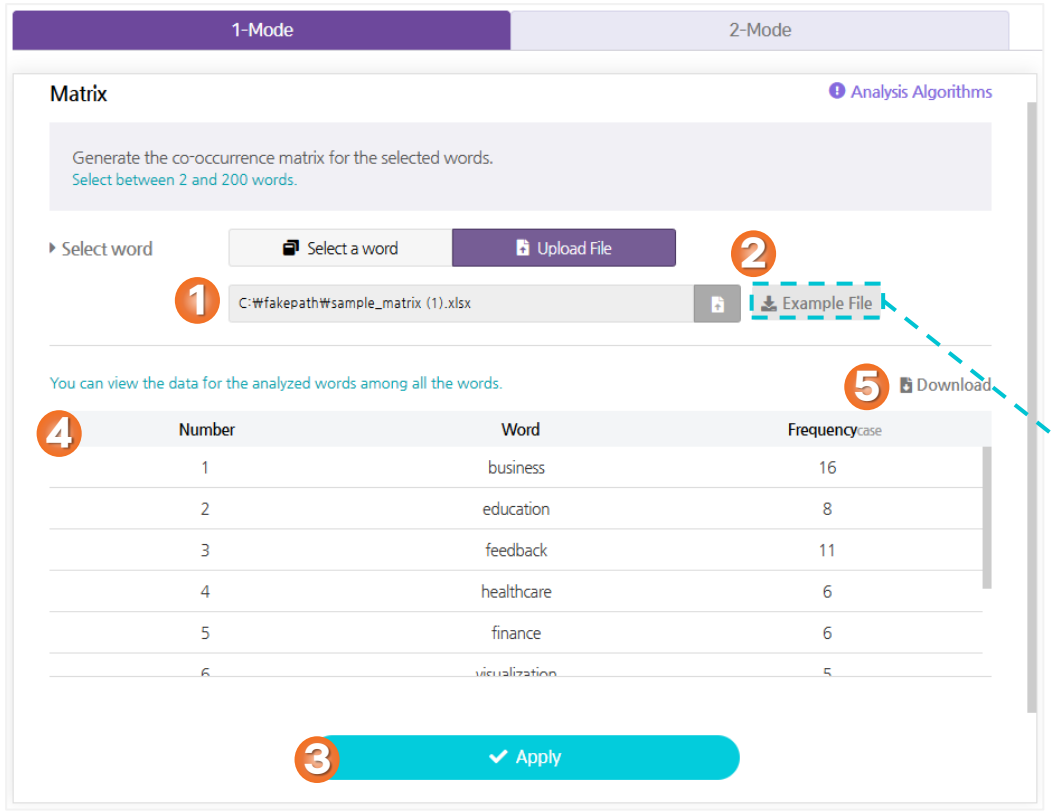
⑤ **Download Analysis Results**
Click [Selected Download] or [Download All] to download word frequencies and percentages into an Excel (.xlsx) file.

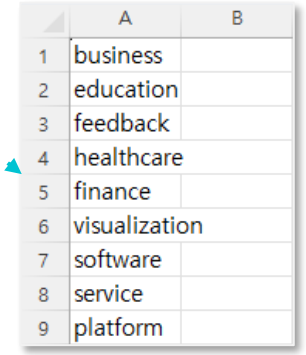
⑥ Apply the selected words to the analysis.

⑦ Display the analysis results.

Matrix

Select words for 1-Mode Matrix Analysis.





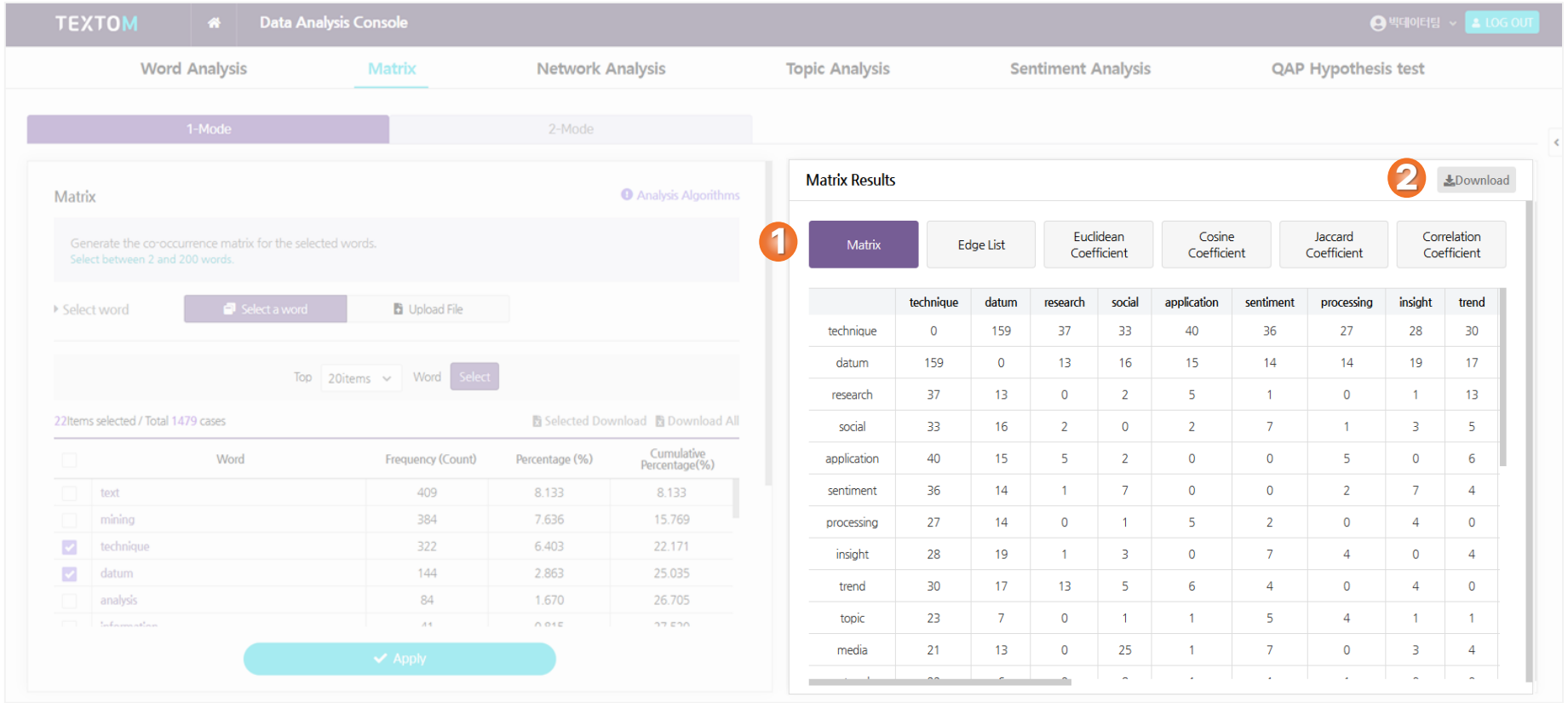
① **File Upload**
Upload an Excel file (.xlsx) containing the matrix word list.

② **Sample File**
Click this button to download a sample format.

③ Apply the selected words to the analysis.

④ **Selected Words Review**
View the list of words applied to the analysis from the entire word set.

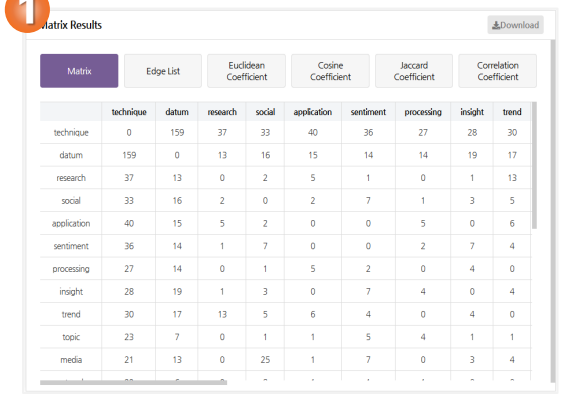
⑤ **Download Word List**
Download the list of words applied to the analysis as an Excel (.xlsx) file.

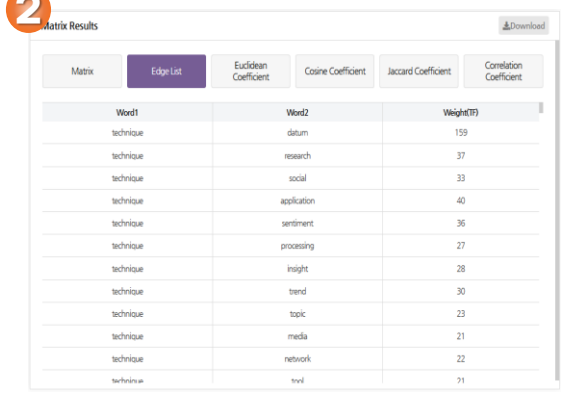


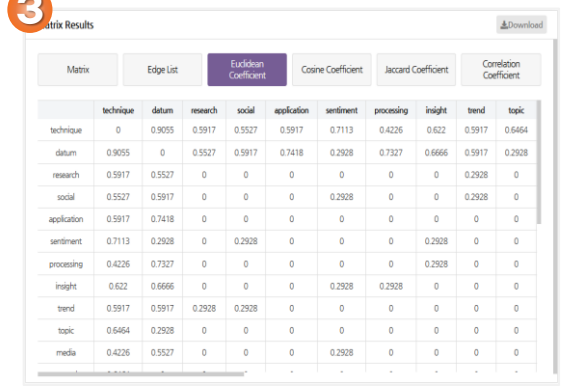
① **Matrix Results**
The analysis results are provided using the following six similarity coefficients. Click on the analysis name to display the corresponding results below.

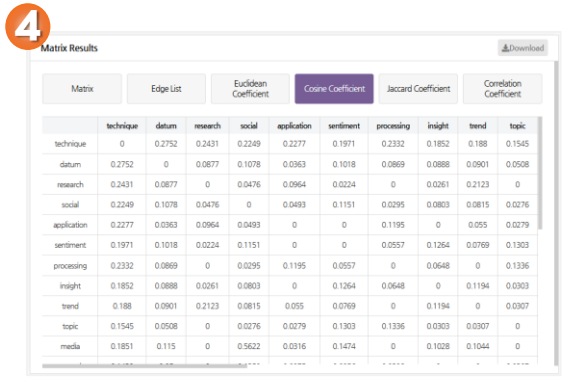
- Matrix
- Edge List
- Euclidean Coefficient
- Cosine Coefficient
- Jaccard Coefficient
- Correlation Coefficient

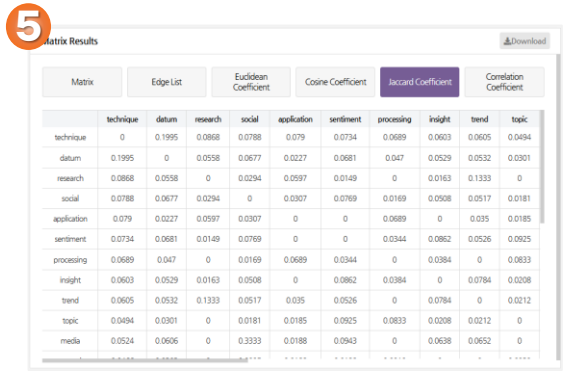
② **Download Analysis Result**
Download the analysis results as an Excel (.xlsx) file.
Tip. Theoption to download all analysis results at once is not available.

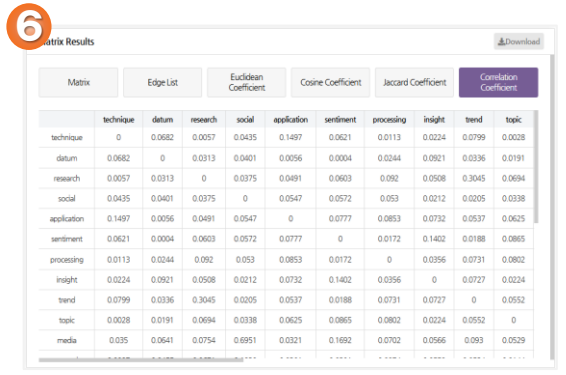












① Matrix
Displays the co-occurrence frequency of selected words within the entire document.

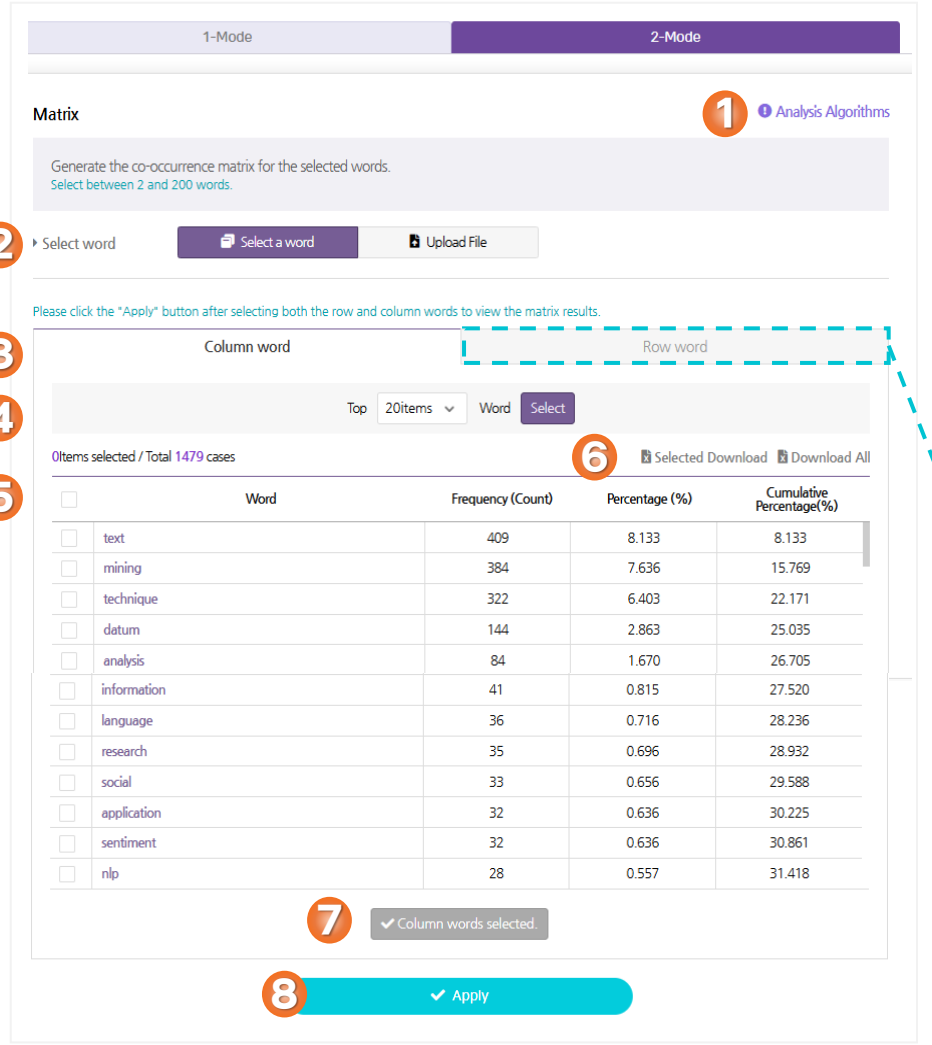
② Edge List
Shows pairs of selected words (word1, word2) and their frequency (Weight) in a list format.
Tip. The option to download all analysis results at once is not available.

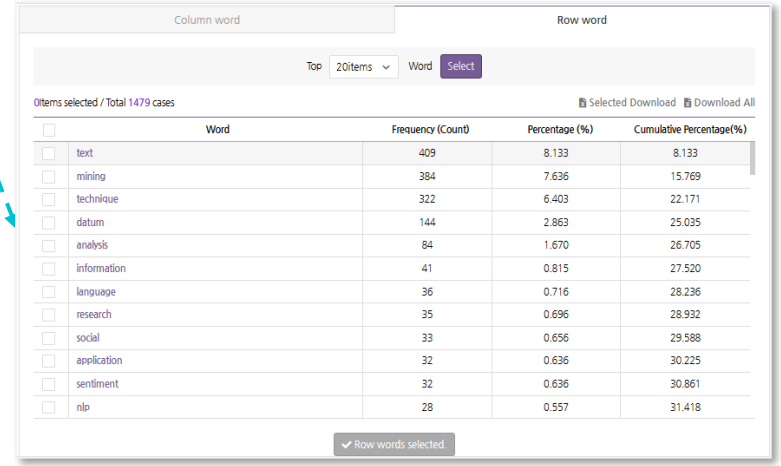
③ Euclidean Coefficient
Measures and provides the similarity between selected words using Euclidean distance.
Tip. The smaller the value between two vectors, the more similar they are.

④ Cosine Coefficient
Calculates and provides the cosine similarity between two words vectors.
Tip. The coefficient ranges from 0 to 1, with values closer to 1 indicating higher similarity between the vectors.

⑤ Jaccard Coefficient
Shows the Jaccard similarity between words.
Tip: The coefficient ranges from 0 to 1, with higher values indicating greater similarity.

⑥ Correlation Coefficient
Indicates the correlation between words.
Tip: Higher values indicate a stronger relationship between the two variables.





① **Selected Algorithm Description**
Click the button to view the description of the algorithm.

② **Set Word Selection Method**
Select 2 to 200 words for analysis. You can choose words directly from the analysis console or upload an Excel file.

③ Select the tab to switch between column words and row words.

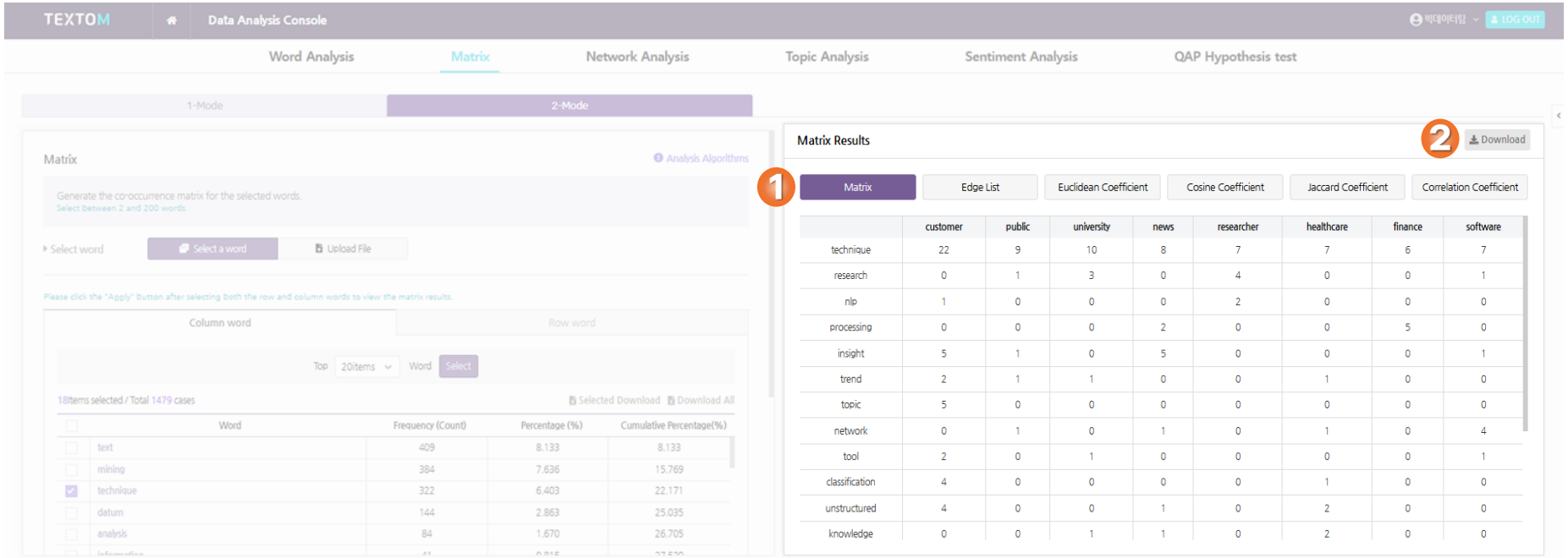
④ **Top Words Auto-Select**
Automatically select the top 20, 50, 75, or 100 words from all words.

⑤ **Word Selection**
Click the top checkbox to automatically select the top 100 words. To manually select words, use the checkbox next to each word.
Tip. Selecting key words related to the analysis topic is crucial. Review the results after the initial selection and adjust as necessary to enhance the accuracy of the analysis.

⑥ **Download Analysis Results**
Click [Selected Download] or [Download All] to download word frequencies and percentages into an Excel (.xlsx) file.

⑦ After selecting the column and row words, click the [Apply] button to save.

⑧ Display the analysis results.



① **Matrix Results**

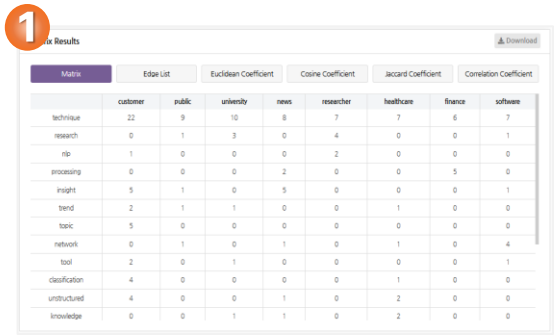
The analysis results are provided using the following six similarity coefficients. Click on the analysis name to display the corresponding results below.

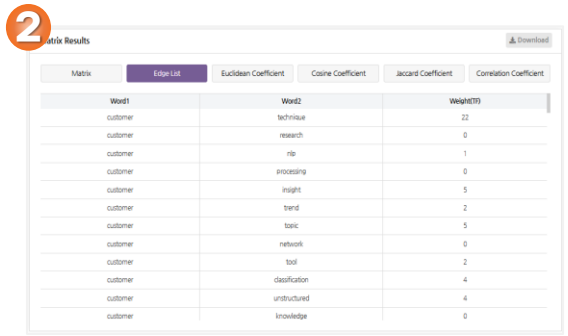
- Matrix
- Edge List
- Euclidean Coefficient
- Cosine Coefficient
- Jaccard Coefficient
- Correlation Coefficient

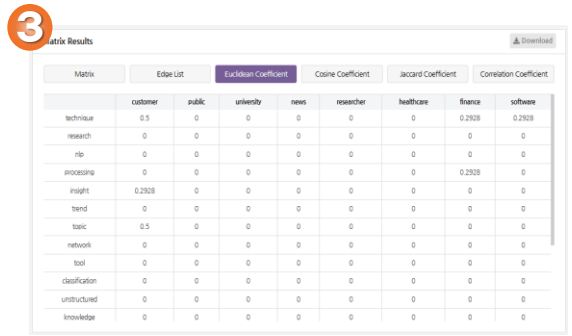
② **Download Analysis Result**

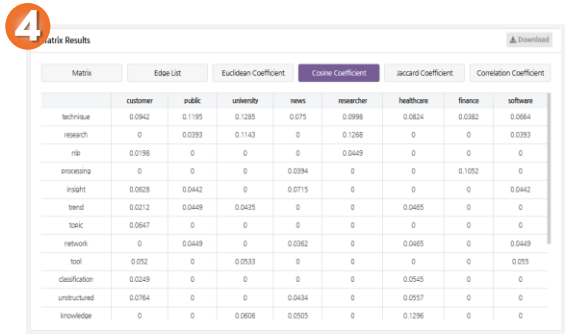
Download the analysis results as an Excel (.xlsx) file.

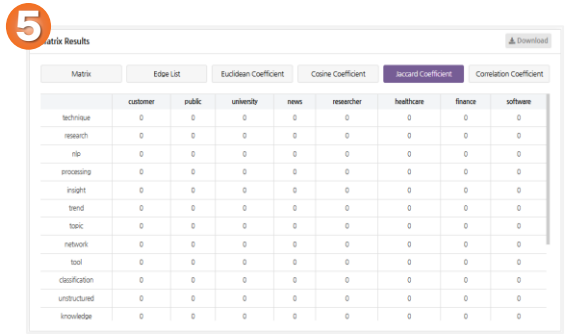
Tip. The option to download all analysis results at once is not available.

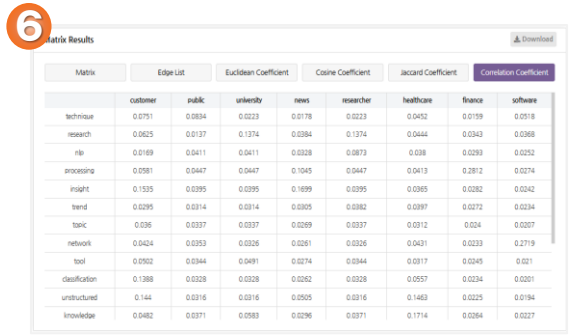












① Matrix
Displays the co-occurrence frequency of selected words within the entire document.

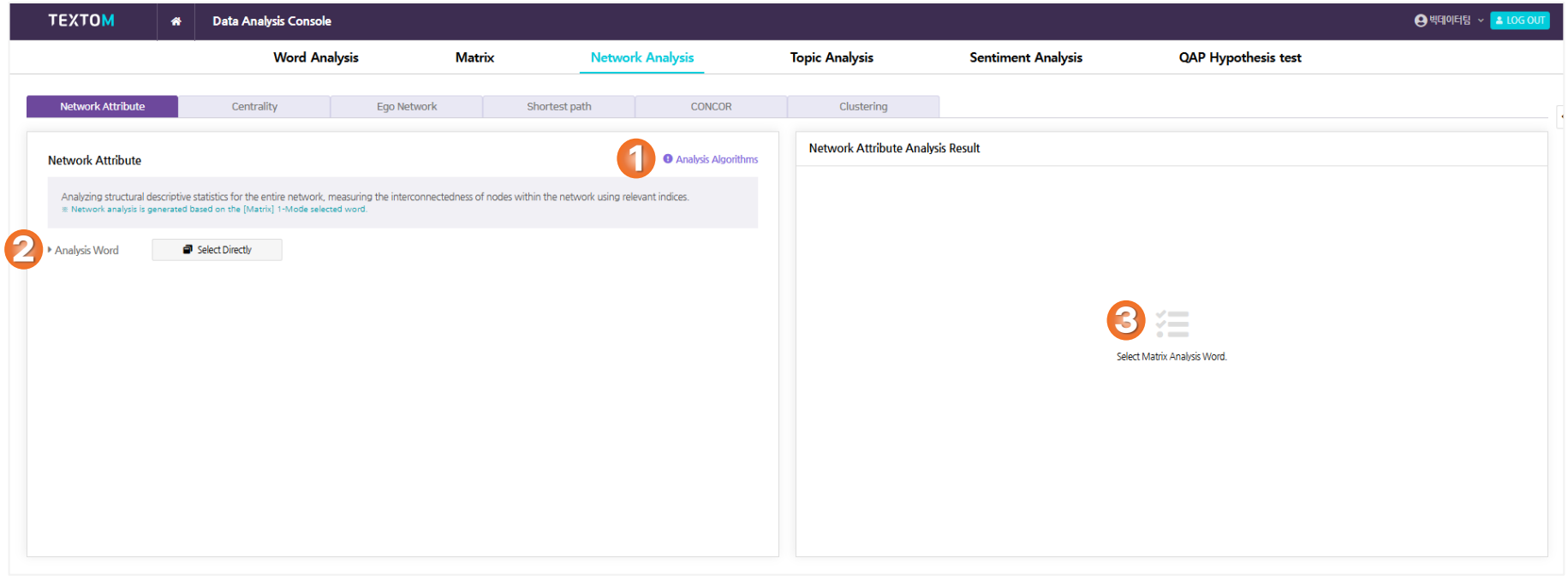
② Edge List
Shows pairs of selected words (word1, word2) and their frequency (Weight) in a list format.
Tip. The option to download all analysis results at once is not available.

③ Euclidean Coefficient
Measures and provides the similarity between selected words using Euclidean distance.
Tip. The smaller the value between two vectors, the more similar they are.

④ Cosine Coefficient
Calculates and provides the cosine similarity between two words vectors.
Tip. The coefficient ranges from 0 to 1, with values closer to 1 indicating higher similarity between the vectors.

⑤ Jaccard Coefficient
Shows the Jaccard similarity between words.
Tip: The coefficient ranges from 0 to 1, with higher values indicating greater similarity.

⑥ Correlation Coefficient
Indicates the correlation between words.
Tip: Higher values indicate a stronger relationship between the two variables.



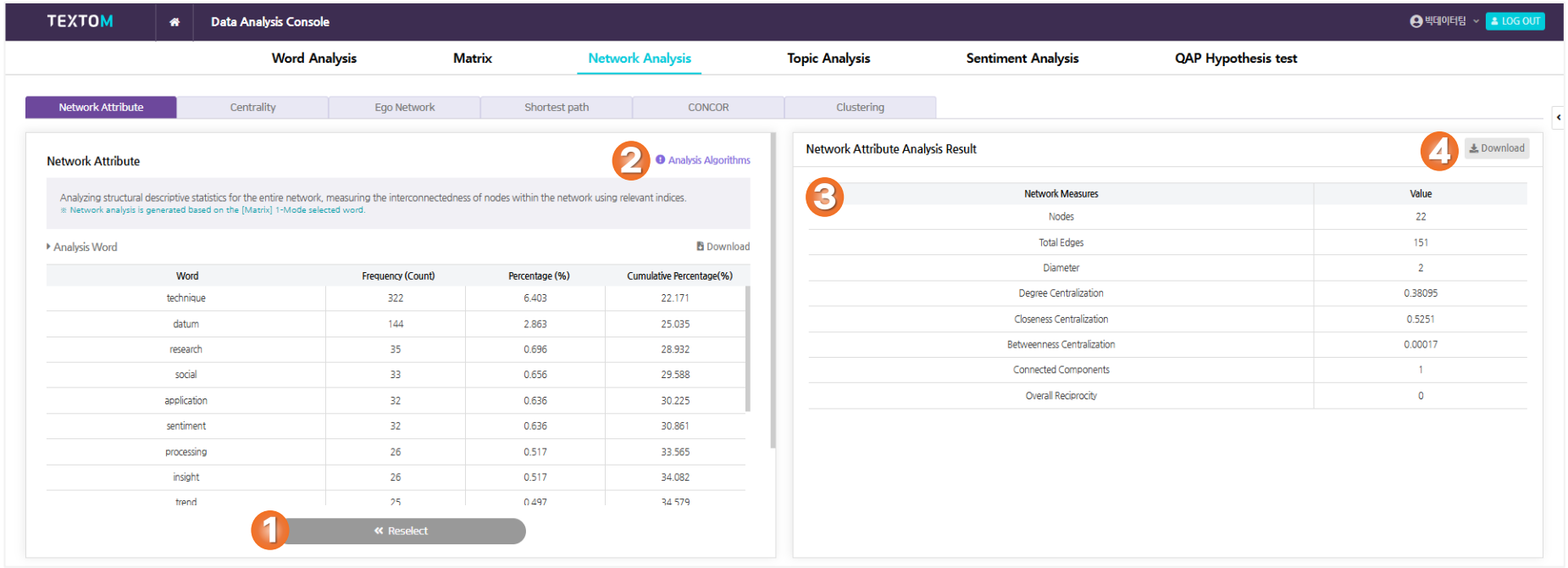
① Selected Algorithm Description
Click the button to view the description of the algorithm.

② Select word
Click [Select Directly] to navigate to the 1-Mode Matrix word selection page and select words for analysis.
Tip. A 1-Mode Matrix analysis must be conducted first.

③ Display the analysis results.

Network Analysis

The result of Analyze Network Attribute.



① **Reselect Words**
Click [Reselect] to navigate to the 1-Mode Matrix word selection page and reselect words for analysis.

② **Download Analysis Words**
Download the selected words as an Excel (.xlsx) file.

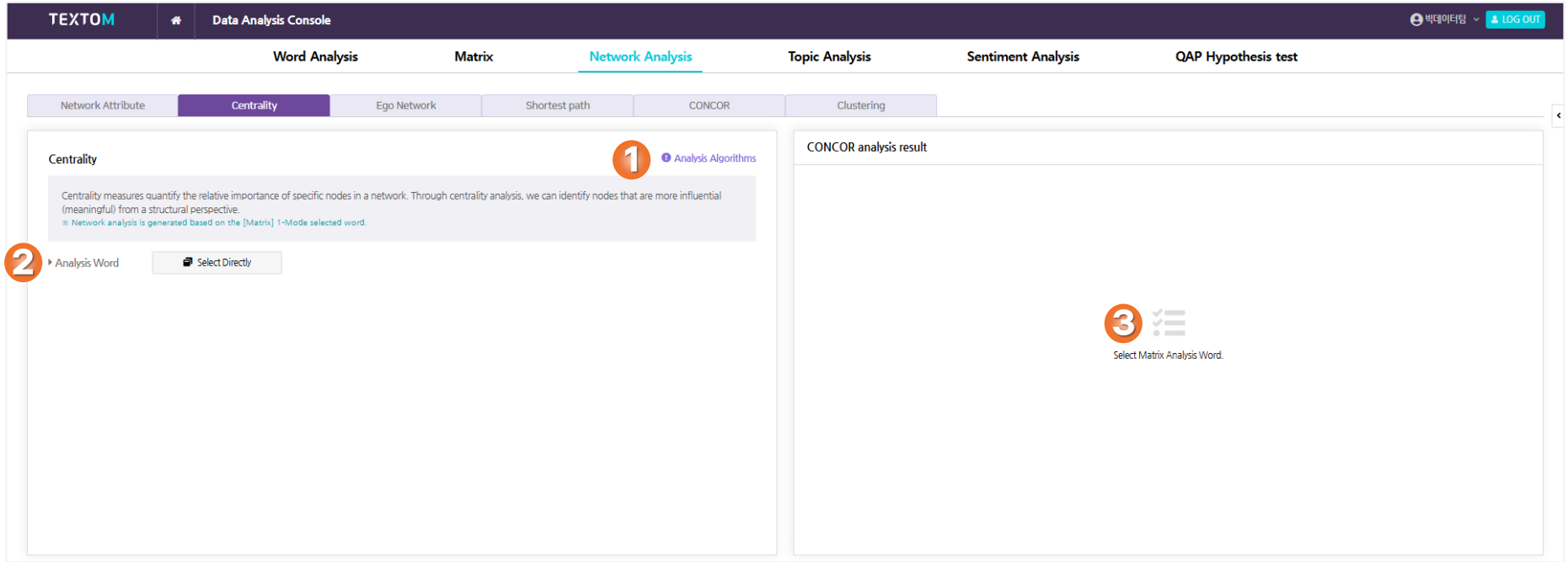
③ **Network Attribute Analysis Results**
Calculate eight types of structural property metrics for the 1-Mode Matrix.

Nodes
Total Edges
Diameter
Degree Centralization
Closeness Centralization
Betweenness Centralization
Connected Components
Overall Reciprocity

④ **Download Analysis Result**
Download the analysis result as an Excel (.xlsx) file.

Network Analysis

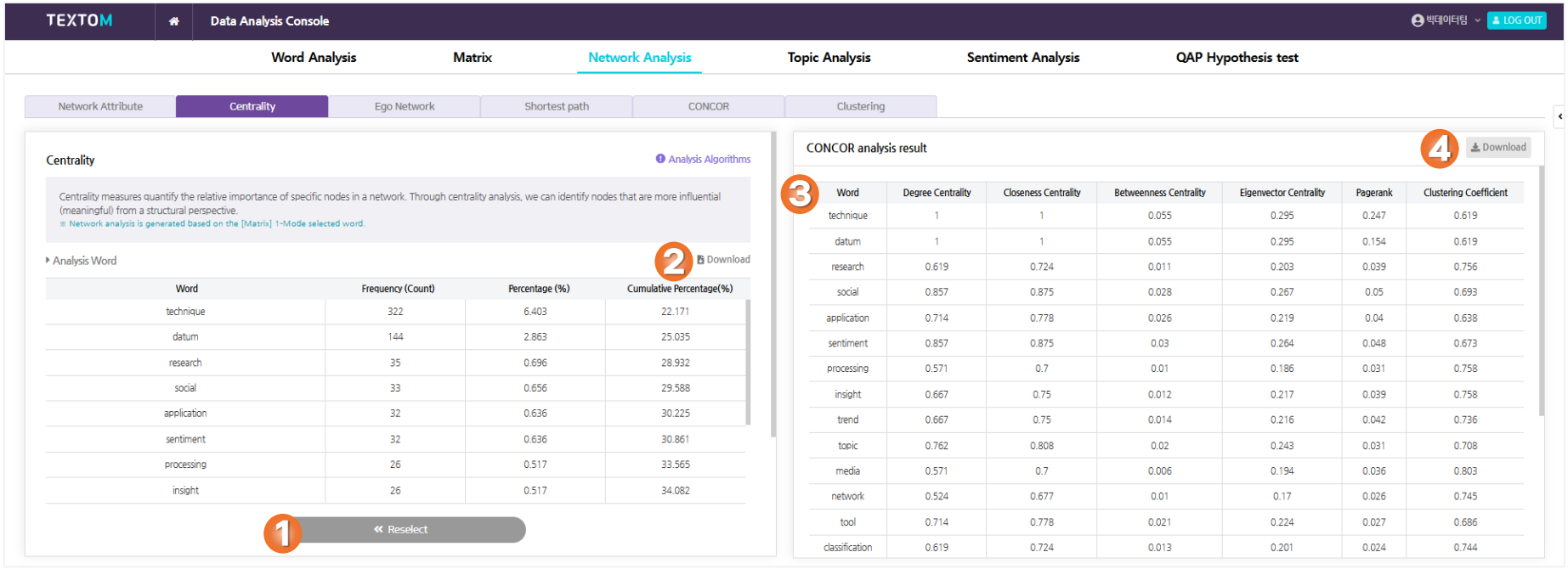
Perform Centrality Analysis.



① Selected Algorithm Description
Click the button to view the description of the algorithm.

② Select word
Click [Select Directly] to navigate to the 1-Mode Matrix word selection page and select words for analysis.
Tip. A 1-Mode Matrix analysis must be conducted first.

③ Display the analysis results.



① **Reselect Words**
Click [Reselect] to navigate to the 1-Mode Matrix word selection page and reselect words for analysis.

② **Download Analysis Words**
Download the selected words as an Excel (.xlsx) file.

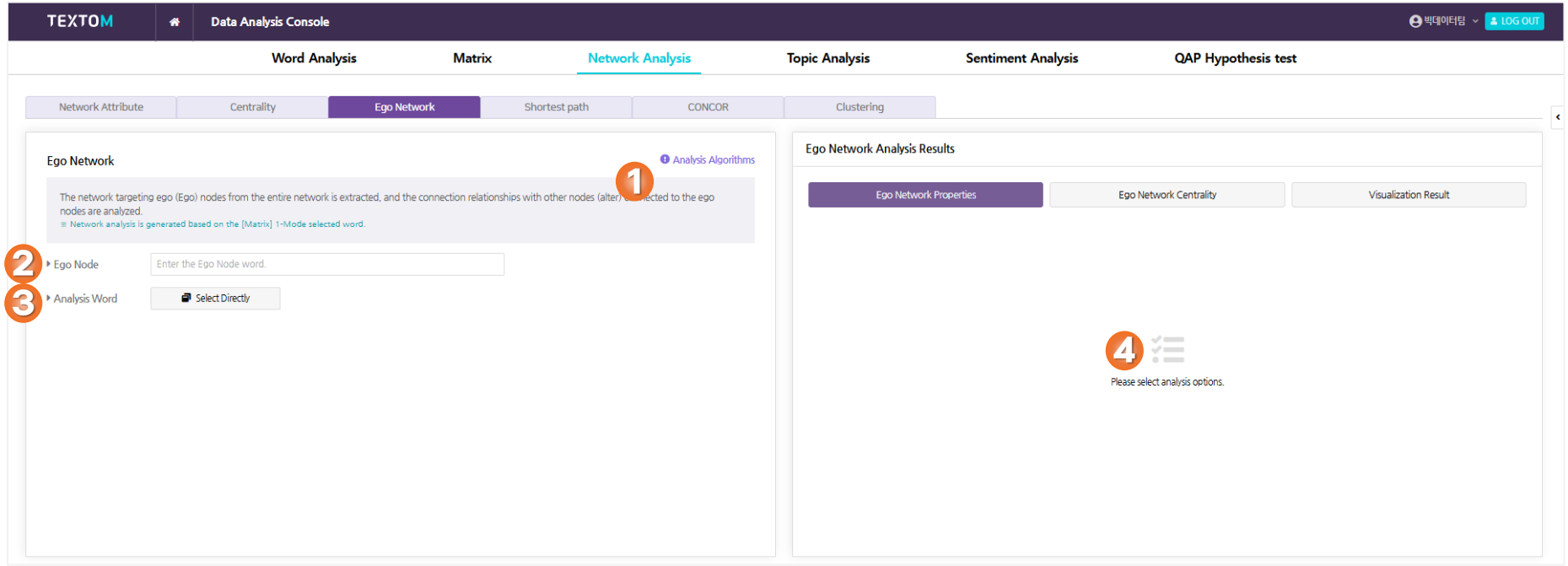
③ **Centrality Analysis Result**
Displays five types of centrality metrics and clustering coefficients. The centrality metrics are as follows.

Degree Centrality
Closeness Centrality
Betweenness Centrality
Eigenvector Centrality
PageRank

④ **Download Analysis Result**
Download the analysis result as an Excel (.xlsx) file.

Network Analysis

Perform Ego Network Analysis.



① Selected Algorithm Description
Click the button to view the description of the algorithm.

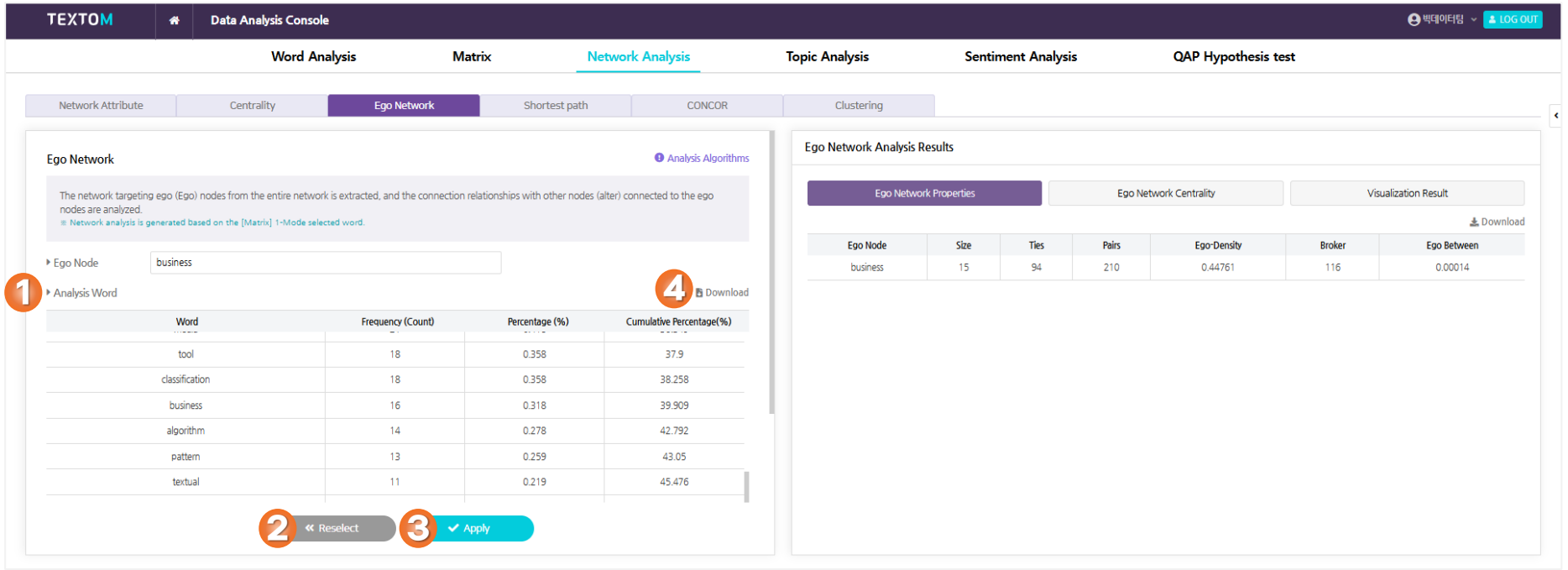
② Enter Ego Node Word
Input the word that will be the focus of the analysis. The network will be extracted based on this central word.

③ Select word
Select the word to analyze in relation to the ego node (word).
Tip. A 1-Mode Matrix analysis must be conducted first.

④ Display the analysis results.

Network Analysis

The results of Ego Network Analysis.

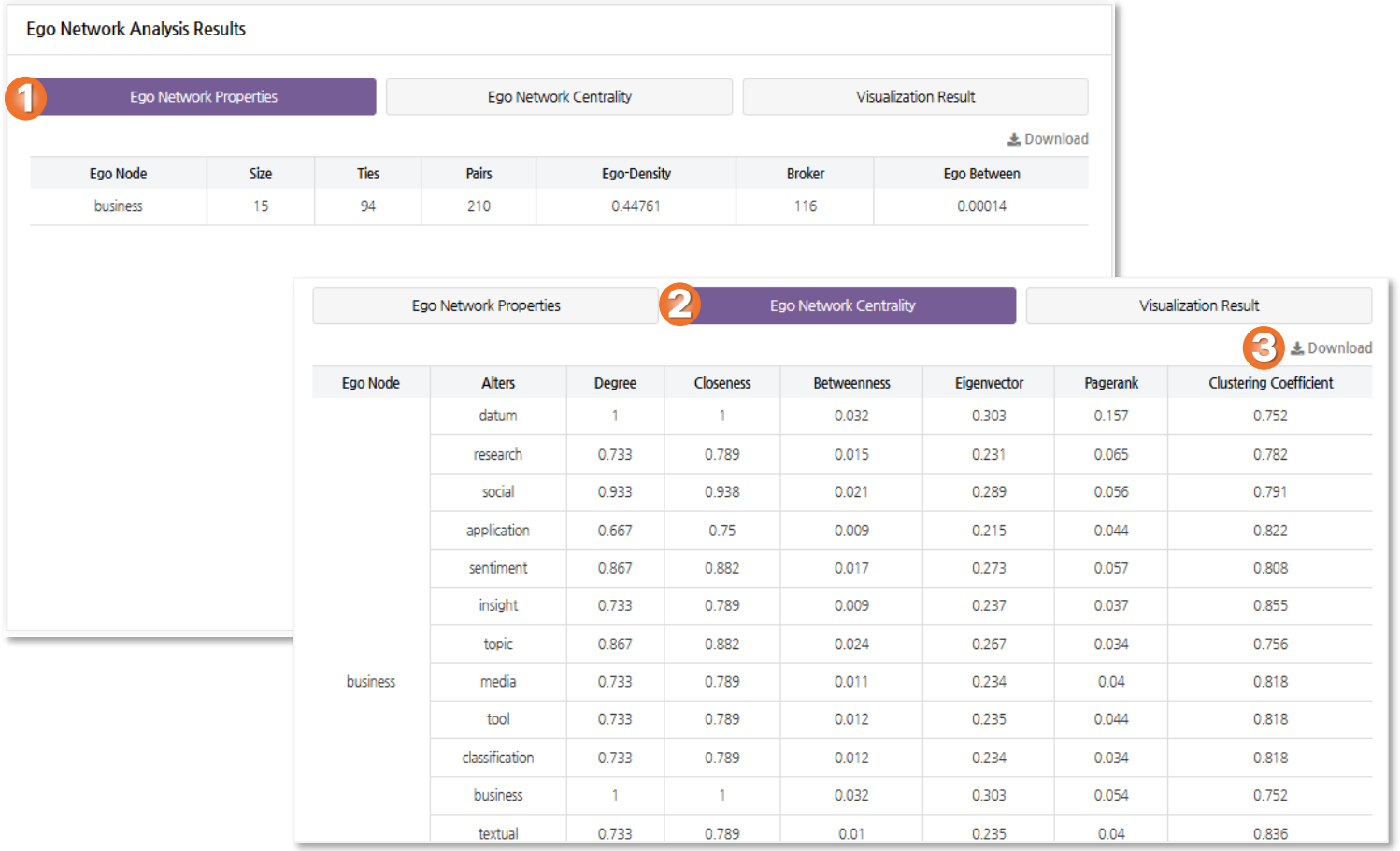


① Review analysis word list.

② Reselect Words
Click [Reselect] to navigate to the 1-Mode Matrix word selection page and reselect
Tip. A 1-Mode Matrix analysis must be conducted first.

③ Apply the words to be analyzed alongside the ego node to the analysis.

④ Download Analysis Words
Download the selected words as an Excel (.xlsx) file.



① **Ego Network Properties**
Provides six types of ego network metrics as analysis results.

Size
Ties
Pairs
Ego-Density
Broker
Ego Betweenness

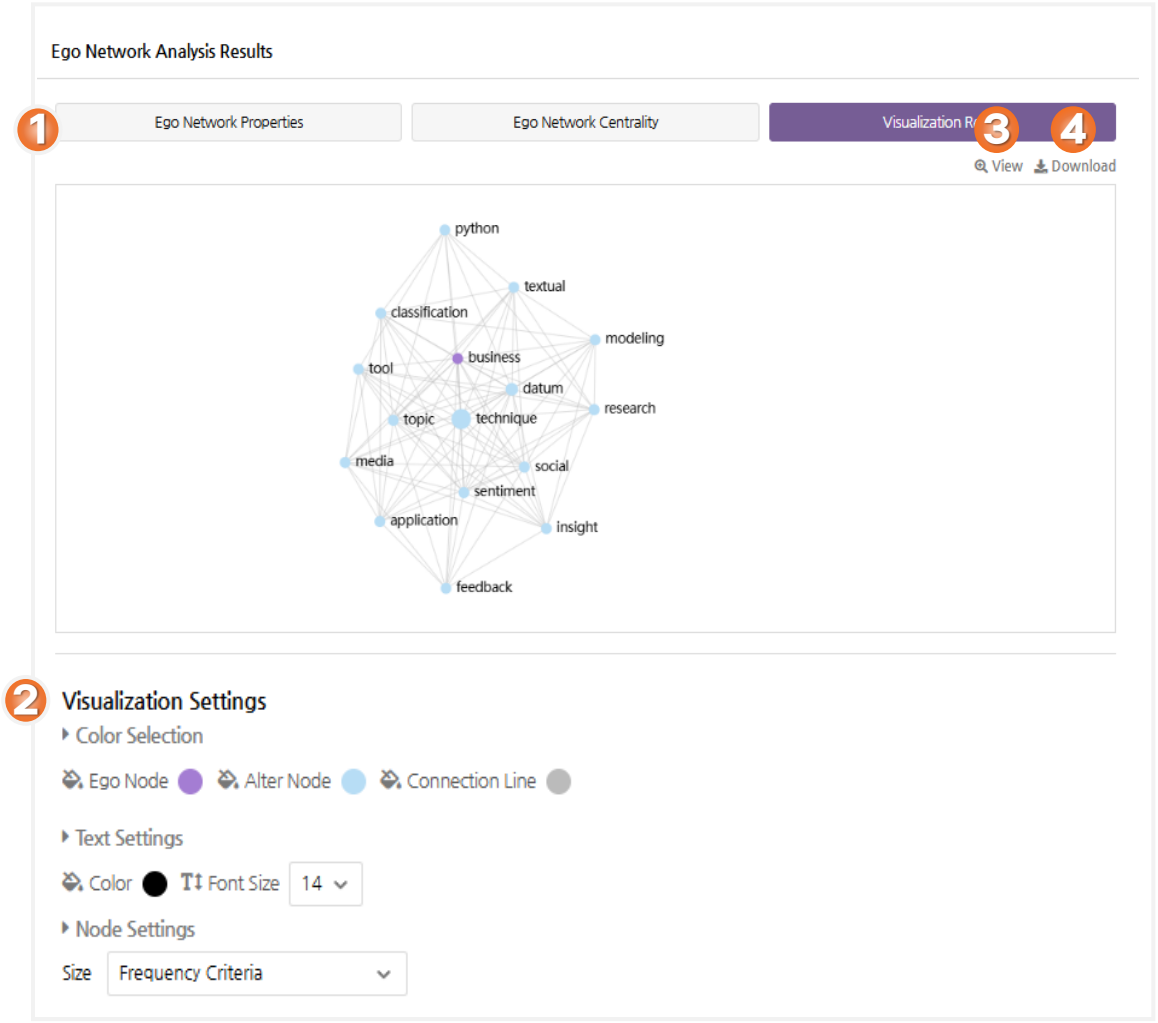
② **Ego Network Centrality**
Displays five types of centrality metrics and the clustering coefficient as analysis results. The centrality metrics are as follows.

Degree Centrality
Closeness Centrality
Betweenness Centrality
Eigenvector Centrality
PageRank

③ **Download Analysis Result**
Download the analysis results as an Excel (.xlsx) file.

Network Analysis

Displays the visualization results of Ego Network.



① Visualization Result

The visualization result of ego network is provided as an undirected network graph.

Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

② Visualization Setting

Users can modify the attributes of the visualization results as desired.

Tip. When set based on word frequency, frequently occurring words are displayed as larger nodes. When set based on betweenness centrality, words that play a more important role in the network are displayed as larger nodes.

③ Zoom In

Click [View] to display an enlarged version of the selected visualization in a pop-up window.

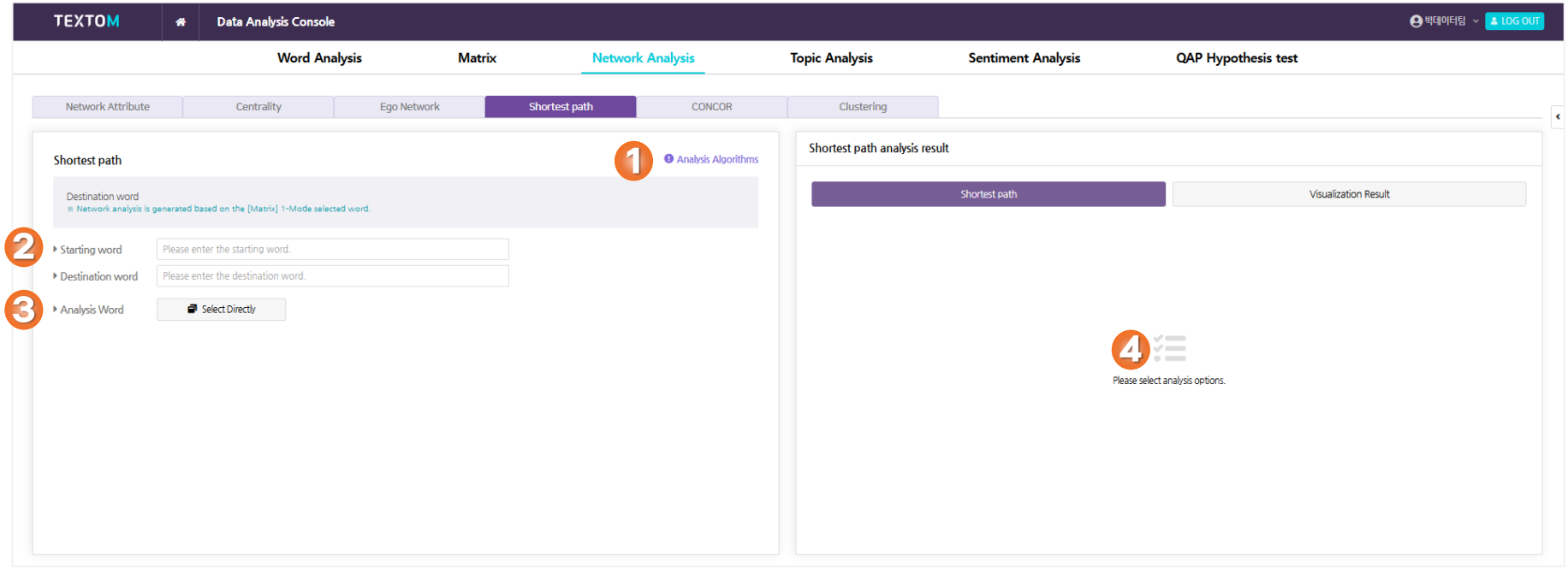
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

④ Download Visualization

Download the visualization result as a PNG file.

Network Analysis

Perform Shortest Path Analysis.



① Selected Algorithm Description
Click the button to view the description of the algorithm.

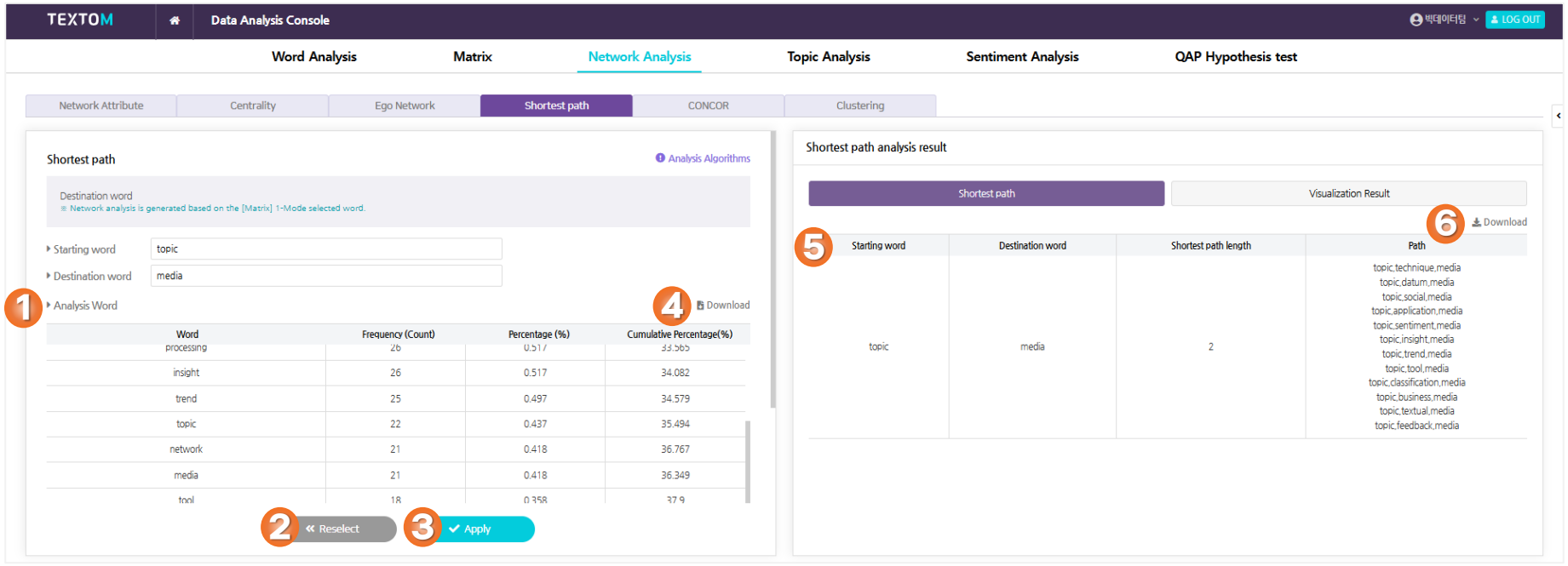
② Enter the initial start(source) word and the destination word to check the path within the network.
Tip. Be sure to enter words that exist within the network.

③ Select word
Click [Select Directly] to navigate to the 1-Mode Matrix word selection page and select words for analysis.
Tip. A 1-Mode Matrix analysis must be conducted first.

④ Display the analysis results.

Network Analysis

The results of Shortest Path Analysis.



① Review analysis word list.

② Reselect Words
Click [Reselect] to navigate to the 1-Mode Matrix word selection page and reselect.
Tip. A 1-Mode Matrix analysis must be conducted first.

③ Apply the selected words to the analysis.

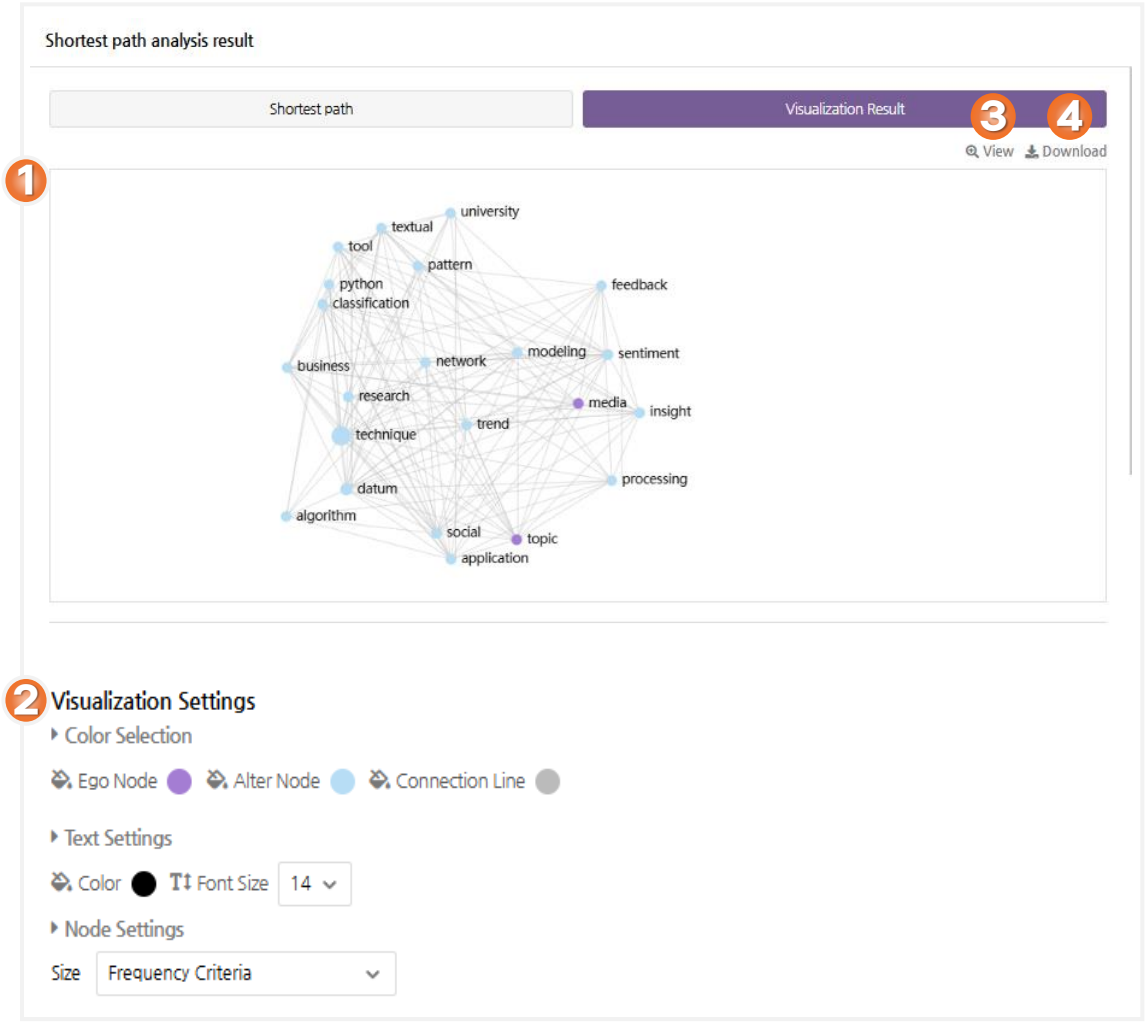
④ Download Analysis Words
Download the selected words as an Excel (.xlsx) file.

⑤ Shortest Path Analysis Results
The result display the shortest path length and detailed paths.

⑥ Download Analysis Result
Download the analysis result as an Excel (.xlsx) file.

Network Analysis

Displays the visualization results of Shortest Path Analysis.



① **Visualization Result**
The results of analysis are presented in a network chart.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

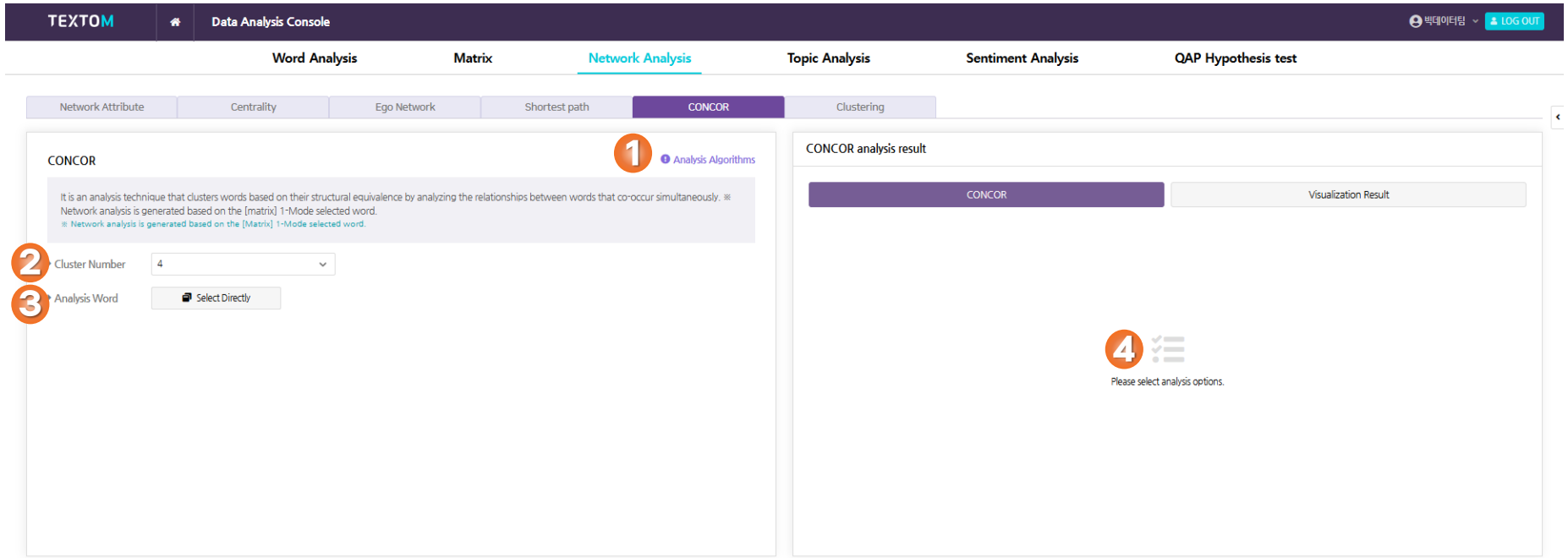
② **Visualization Setting**
Users can modify the attributes of the visualization results as desired.
Tip. When set based on word frequency, frequently occurring words are displayed as larger nodes. When set based on betweenness centrality, words that play a more important role in the network are displayed as larger nodes.

③ **Zoom In**
Click [View] to display an enlarged version of the selected visualization in a pop-up window.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

④ **Download Visualization**
Download the visualization result as a PNG file.

Network Analysis

Perform CONCOR Analysis.

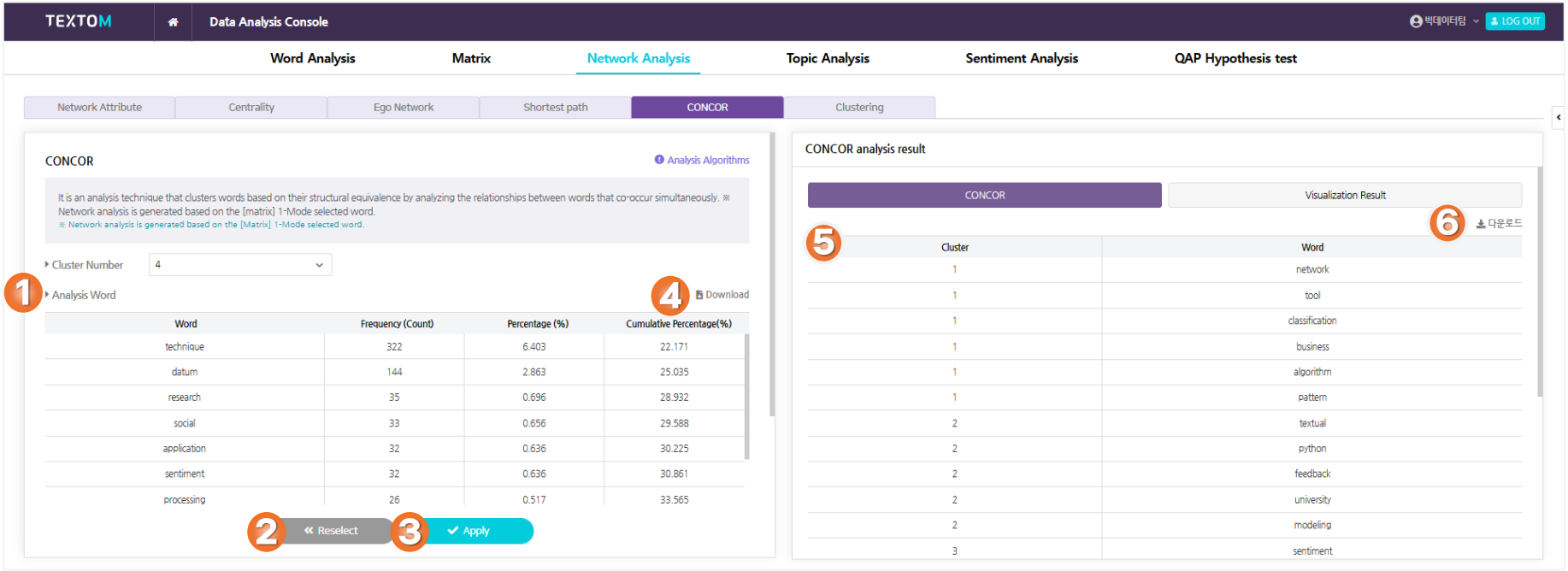


① **Selected Algorithm Description**
Click the button to view the description of the algorithm.

② Select Number Of Cluster.

③ **Select word**
Click [Select Directly] to navigate to the 1-Mode Matrix word selection page and select words for analysis.
Tip. A 1-Mode Matrix analysis must be conducted first.

④ Display the analysis results.



① Review analysis word list.

② **Reselect Words**
Click [Reselect] to navigate to the 1-Mode Matrix word selection page and reselect.
Tip. A 1-Mode Matrix analysis must be conducted first.

③ Apply the selected words to the analysis.

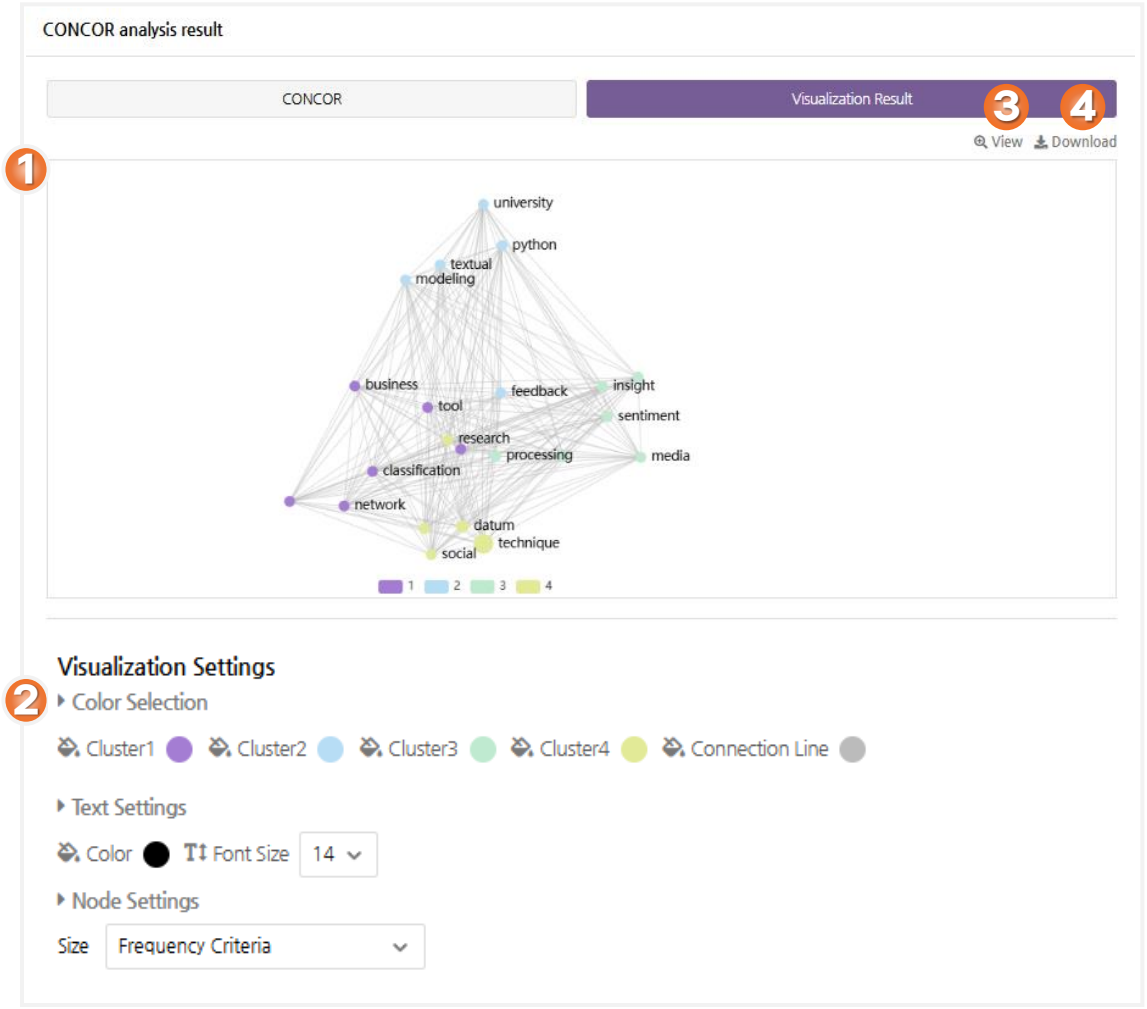
④ **Download Analysis Words**
Download the selected words as an Excel (.xlsx) file.

⑤ **Shortest Path Analysis Results**
The results display the words and the groups to which they belong.

⑥ **Download Analysis Result**
Download the analysis result as an Excel (.xlsx) file.

Network Analysis

Displays the visualization results of CONCOR Analysis.

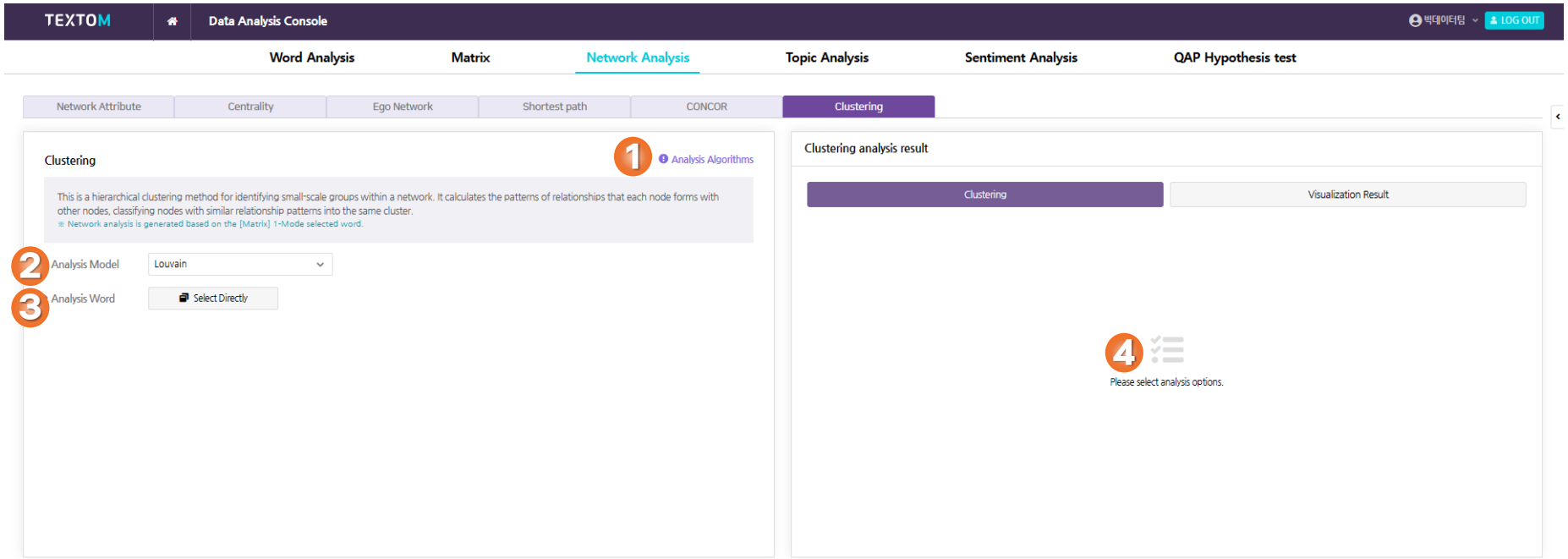


① **Visualization Result**
The results of analysis are presented in a network chart.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

② **Visualization Setting**
Users can modify the attributes of the visualization results as desired.
Tip. When set based on word frequency, frequently occurring words are displayed as larger nodes. When set based on betweenness centrality, words that play a more important role in the network are displayed as larger nodes.

③ **Zoom In**
Click [View] to display an enlarged version of the selected visualization in a pop-up window.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

④ **Download Visualization**
Download the visualization result as a PNG file.



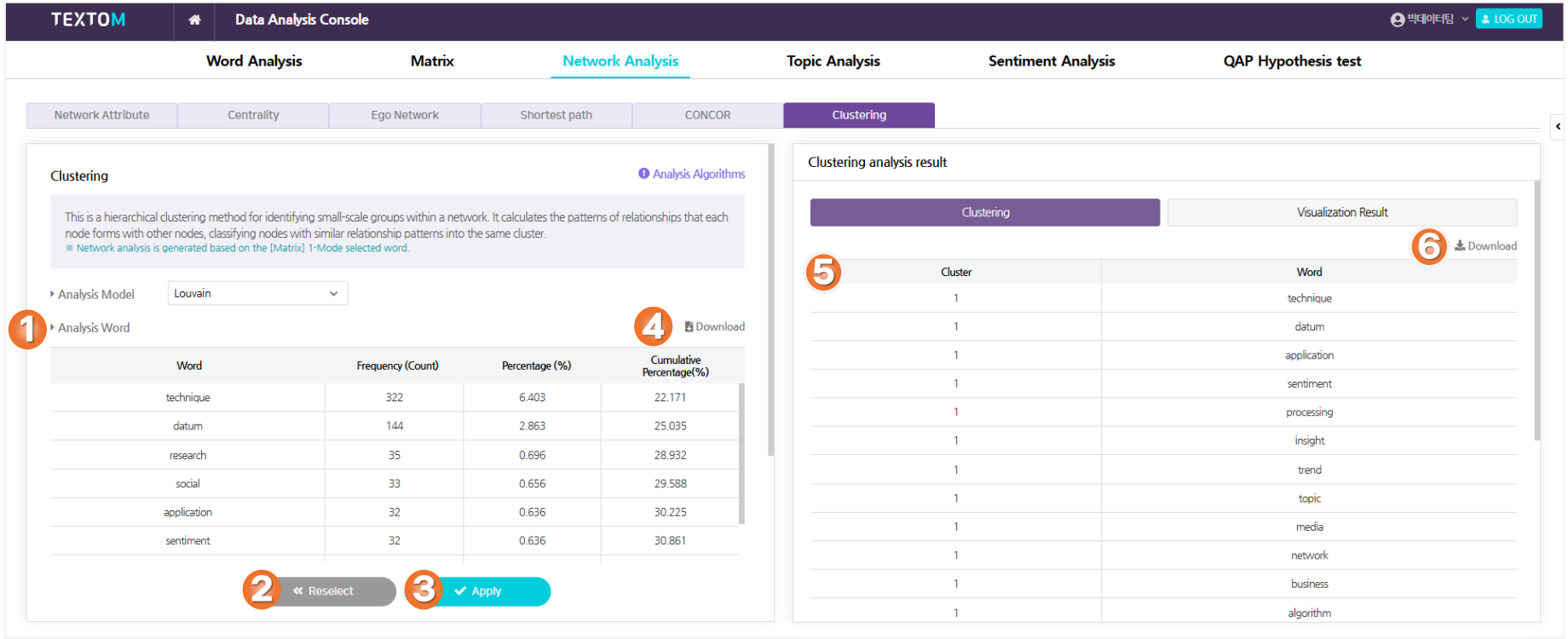
① **Selected Algorithm Description**
Click the button to view the description of the algorithm.

② **Select Analysis Model**
Clustering analysis offers the following four algorithms:
- Louvain
- Leiden
- Girvan Newman (GN)
- Clauset-Newman-Moore

Tip. For quick analysis: Louvain
For high accuracy: Leiden
For simple network structures: GN
For hierarchical structures: Clauset-Newman-Moore

③ **Select word**
Click [Select Directly] to navigate to the 1-Mode Matrix word selection page and select words for analysis.
Tip. A 1-Mode Matrix analysis must be conducted first.

④ Display the analysis results.



① Review analysis word list.

② Reselect Words
Click [Reselect] to navigate to the 1-Mode Matrix word selection page and reselect.
Tip. A 1-Mode Matrix analysis must be conducted first.

③ Apply the selected words to the analysis.

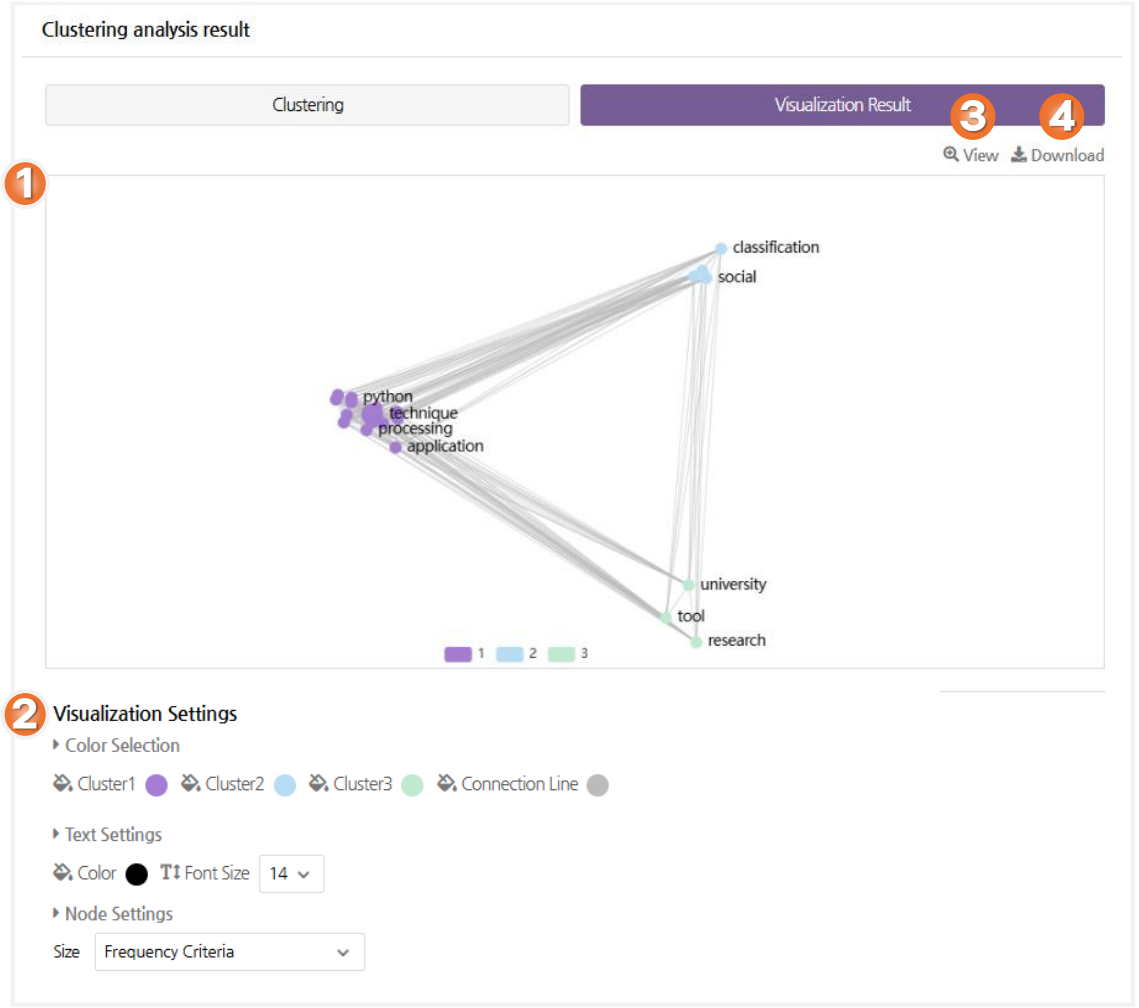
④ Download Analysis Words
Download the selected words as an Excel (.xlsx) file.

⑤ Clustering Analysis Results
The results display the words and the clusters to which they belong.

⑥ Download Analysis Result
Download the analysis result as an Excel (.xlsx) file.

Network Analysis

Displays the visualization results of Clustering Analysis.



① Visualization Result
The results of analysis are presented in a network chart.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

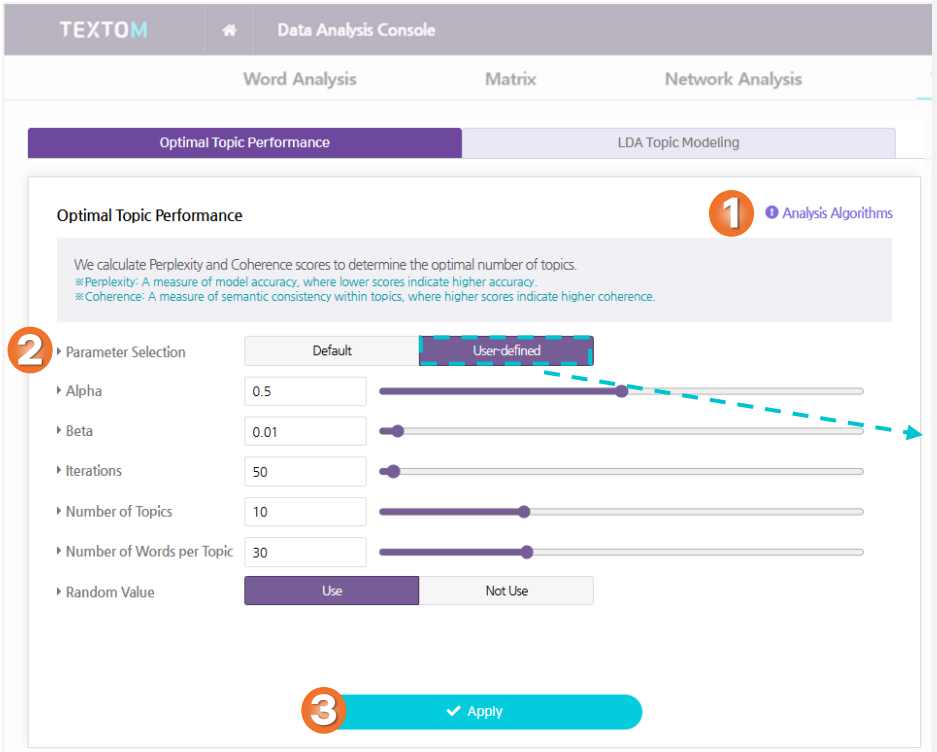
② Visualization Setting
Users can modify the attributes of the visualization results as desired.
Tip. When set based on word frequency, frequently occurring words are displayed as larger nodes. When set based on betweenness centrality, words that play a more important role in the network are displayed as larger nodes.

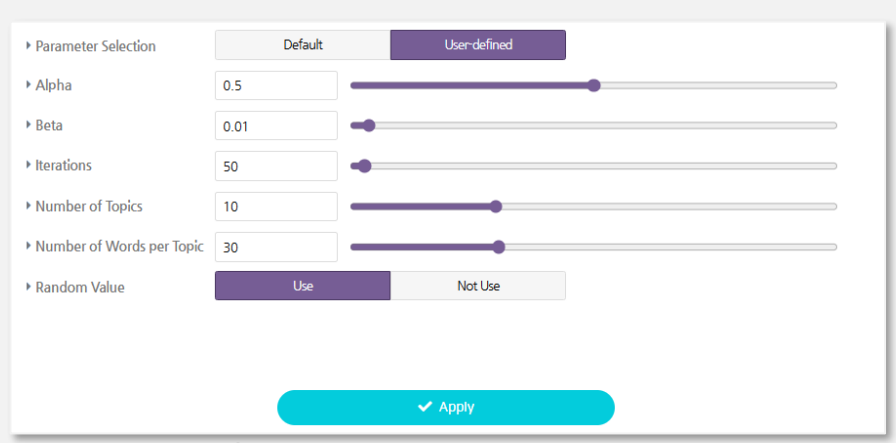
③ Zoom In
Click [View] to display an enlarged version of the selected visualization in a pop-up window.
Tip. You can zoom in and out of the visualization by scrolling the mouse wheel up and down.

④ Download Visualization
Download the visualization result as a PNG file.

Topic Analysis

Check Optimal Topic Performance before Conducting LDA Analysis.





Alpha	A hyperparameter that adjusts the distribution of topics in documents
Beta	A hyperparameter that adjusts the distribution of words in topics
Iterations	The number of repetitions during model training
Number of Topics	The number of groups you want to create
Number of Words per Topic	The number of words to be included in each group
Random Seed	Proceed with random assignments (however, it is difficult to reproduce or verify the results of the experiment)

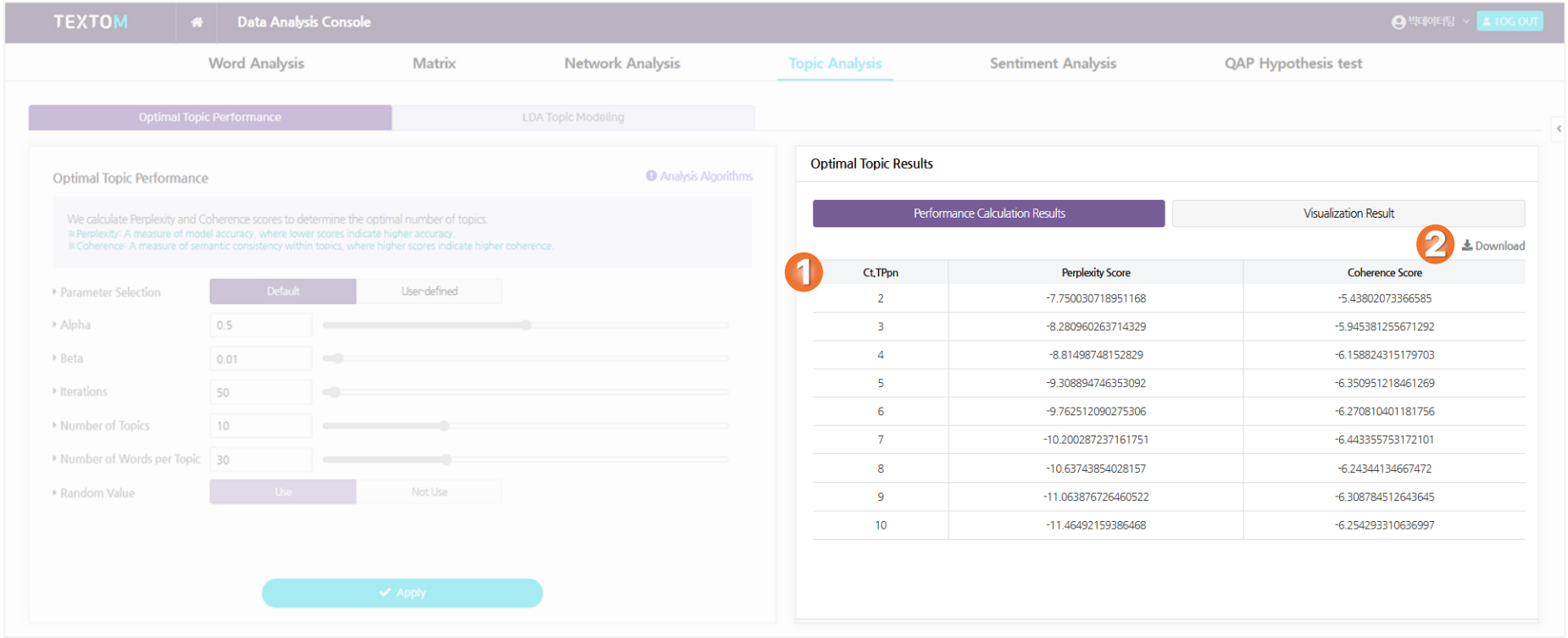
① Selected Algorithm Description
Click the button to view the description of the algorithm.

② LDA Analysis Parameters
Before performing LDA analysis, you can set the parameters to their default values or choose custom settings to achieve optimal topic performance.
Tip. It is recommended to use the default values. After reviewing the topic analysis results, adjust the parameters through custom settings if necessary.

③ Apply the selected options to proceed with the analysis.

Topic Analysis

Display the Optimal Topic Results.



① **Topic Performance Analysis Results**
The following metrics are calculated as topic performance results. Generally, a lower Perplexity value and a higher Coherence value indicate the optimal number of topics.

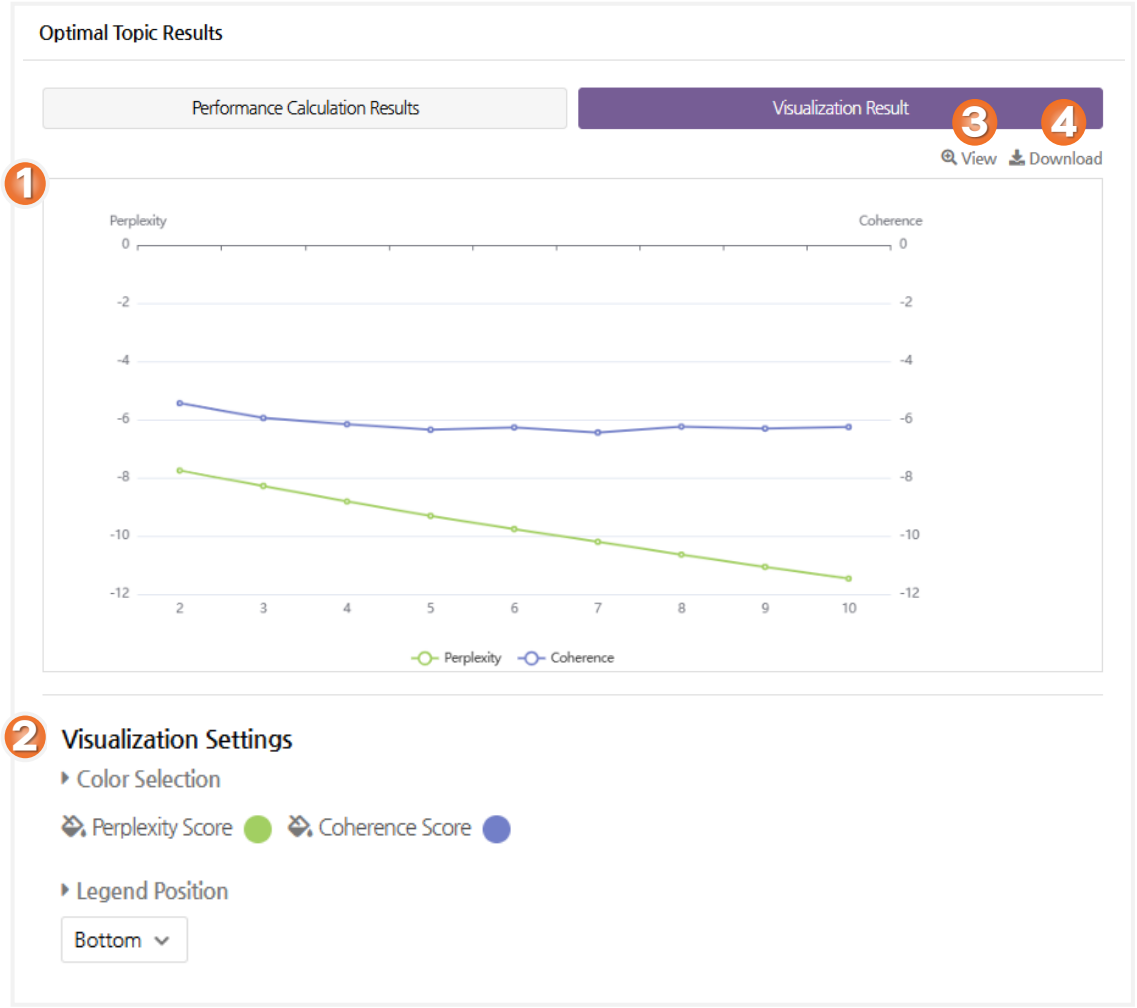
- Perplexity Score
  A metric for the predictive performance of the topic model; lower values indicate higher predictive accuracy.
- Coherence Score
  A metric for evaluating the semantic consistency of words within a topic; higher values indicate more coherent topics.

Tip. After determining the optimal number of topics based on the topic performance analysis results, reset the LDA parameters to this number of topics and proceed with the analysis.

② **Download Analysis Result**
Download the analysis result as an Excel (.xlsx) file.

Network Analysis

Displays the visualization results of Clustering Analysis.

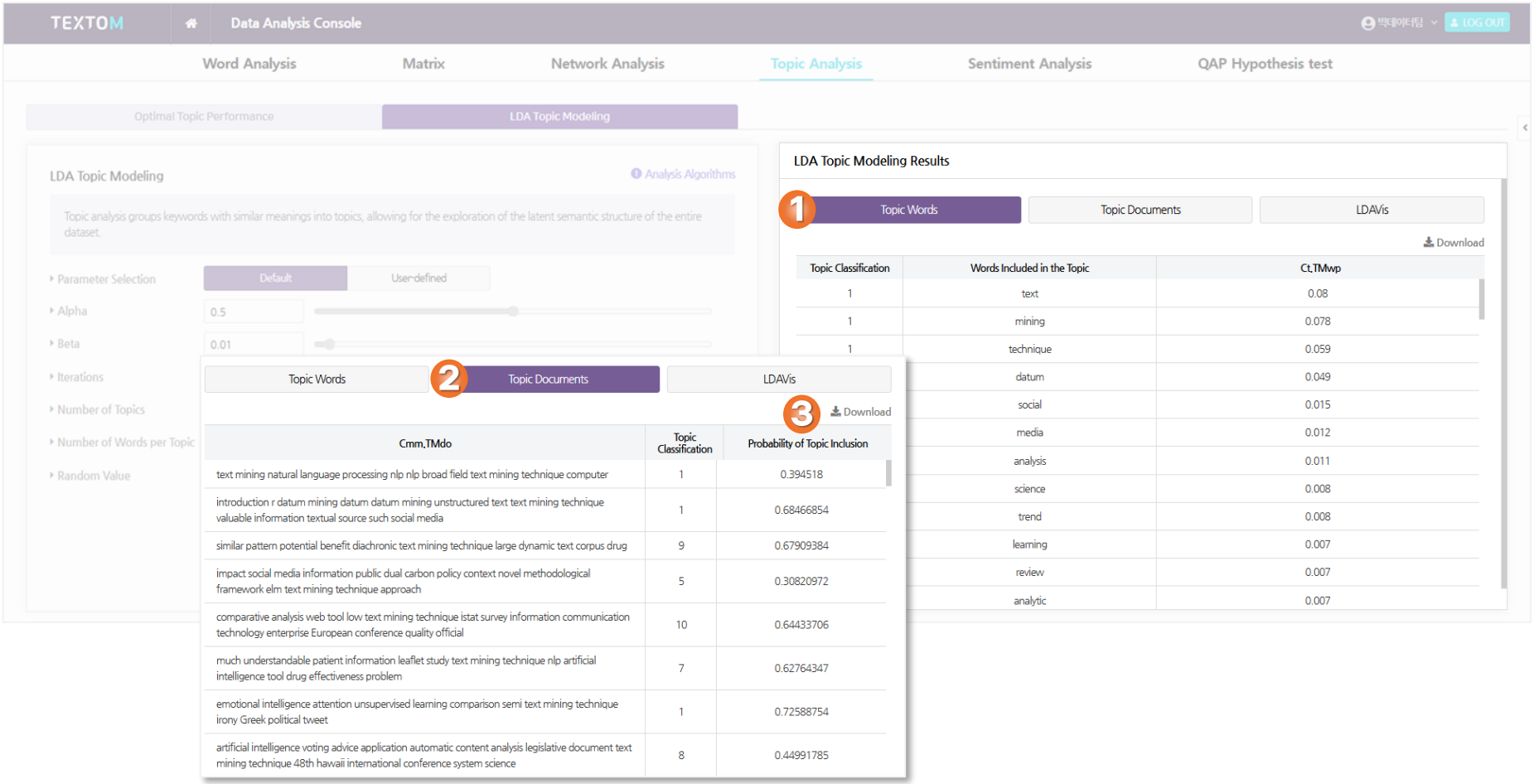


① Visualization Result
The results of analysis are presented in a line chart.

② Visualization Setting
Users can modify the attributes of the visualization results as desired.

③ Zoom In
Click [View] to display an enlarged version of the selected visualization in a pop-up window.

④ Download Visualization
Download the visualization result as a PNG file.



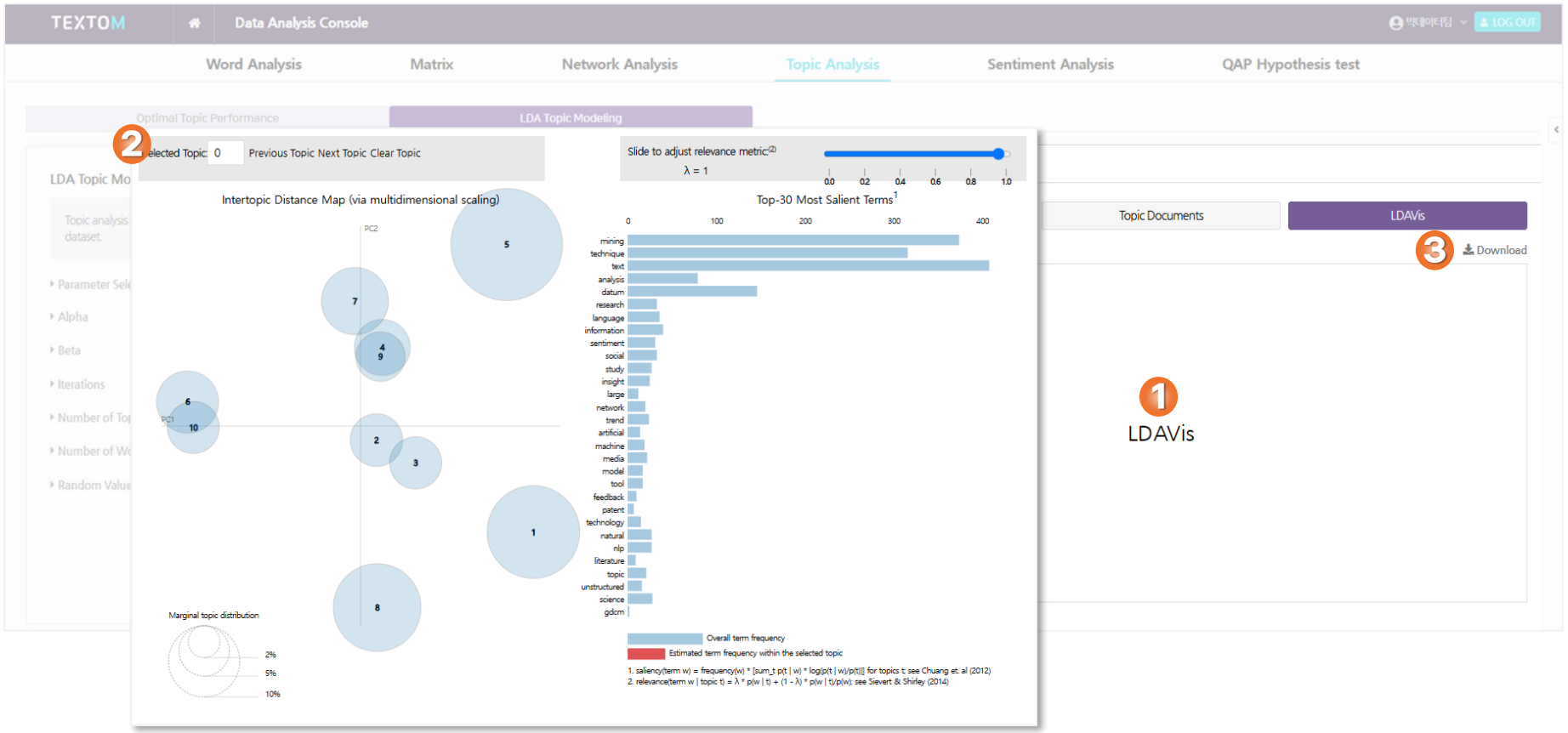
① **Topic Word Analysis Results**
Calculate the extracted topics and the topic probability(%) of words from the topic word analysis.
Tip. Word topic probability indicates the likelihood of a word being assigned to a topic.

② **Topic Document Analysis Results**
Calculate the original texts assigned to the extracted topics and their topic probability(%) from the topic document analysis.

③ **Download Analysis Result**
Download the analysis result as an Excel (.xlsx) file.

Topic Analysis

The visualization results of LDA Topic Modeling.



① Click this text to open the visualization in a pop-up.

② Visualization Results

Use clicks and cursor movement to view the word distribution for each topic and the similarities between topics.

- **Circle Size** : Represents the word distribution of the topic.
- **Circle Distance** : Indicates the similarity between topics.
- **λ** : Adjusts the distinctiveness between topics. Closer to 1 shows words that frequently appear in one topic, while closer to 0 shows words that do not frequently appear in other topics.

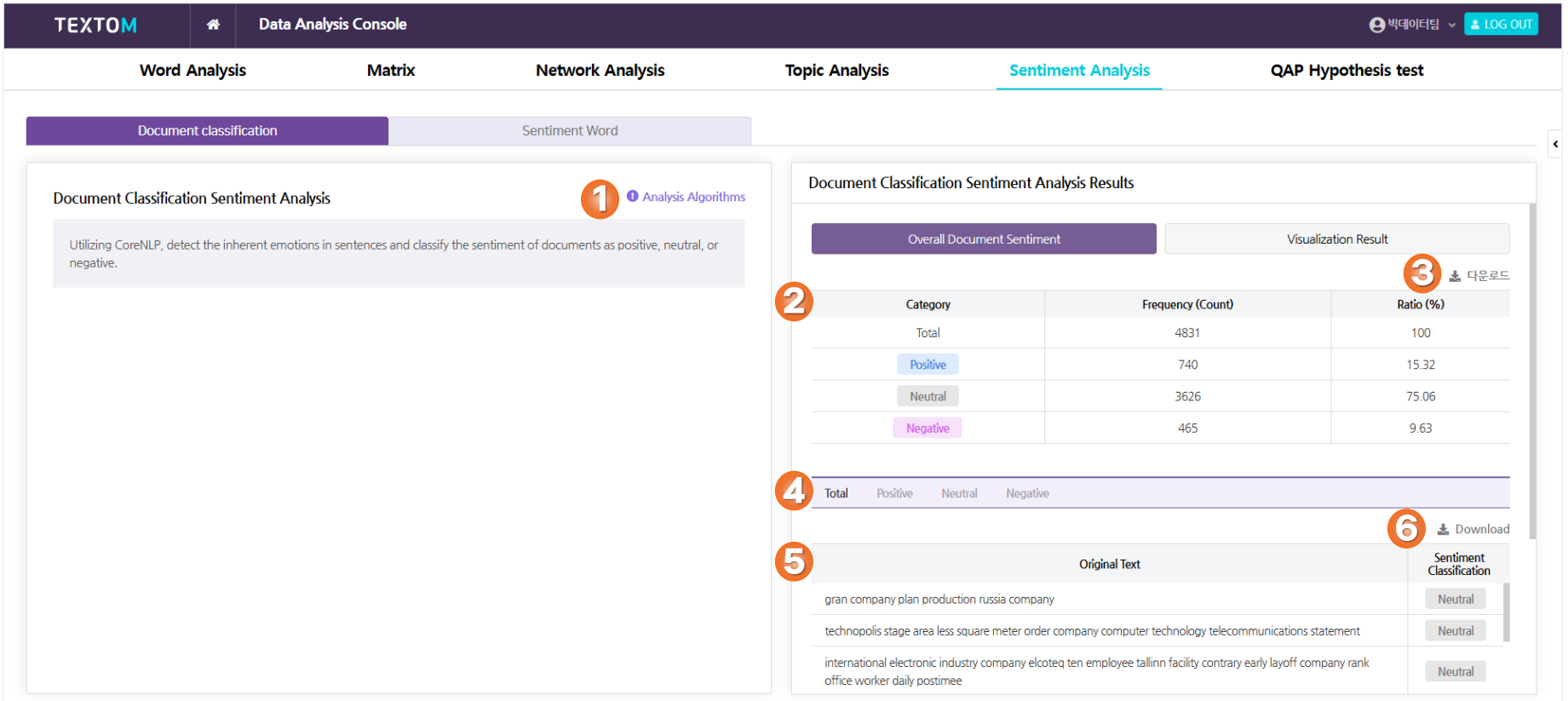
Tip. The visualization shows the top 30 keywords for each topic, so it is recommended to carefully refine words (e.g., single characters, numbers, etc.).

② Download Analysis Result

Download the analysis result as an HTML file.

Sentiment Analysis

Perform Document Classification Sentiment Analysis.



① Selected Algorithm Description
Click the button to view the description of the algorithm.

② Summary of Overall Classification Result
Generate a summary of the frequency and percentage(%) of total/positive/neutral/negative documents from the comprehensive sentiment analysis results.

③ Download the comprehensive sentiment analysis results as an Excel (.xlsx) file.

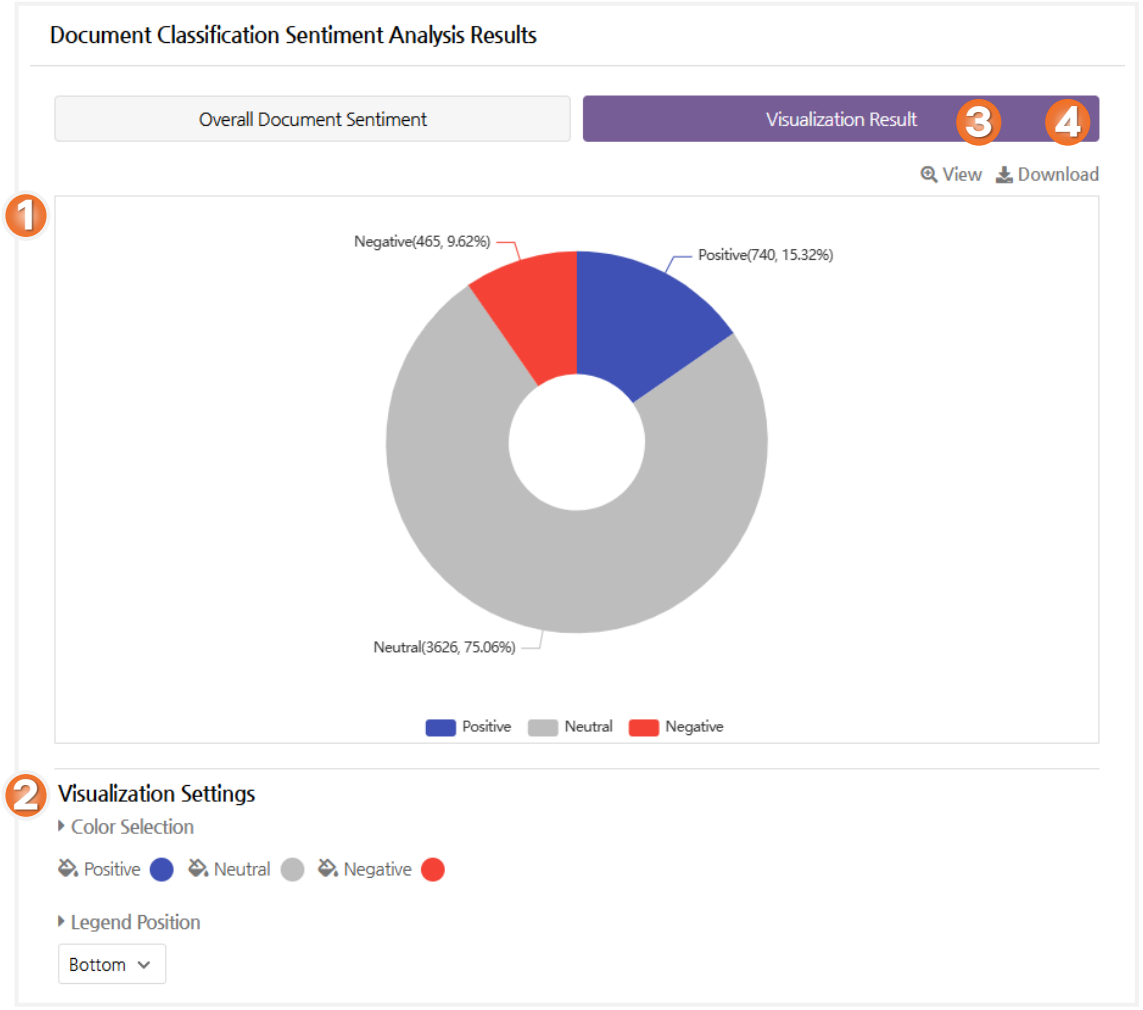
④ Sentiment Classification Tab
Click the Sentiment tab to view the classification results of original texts by sentiment.

⑤ View Sentiment Classification Results
Check the classification results of original texts by sentiment.

⑥ Download the Sentiment Classification result as an Excel (.xlsx) file.

Sentiment Analysis

Displays the visualization results of Document Classification Sentiment Analysis.

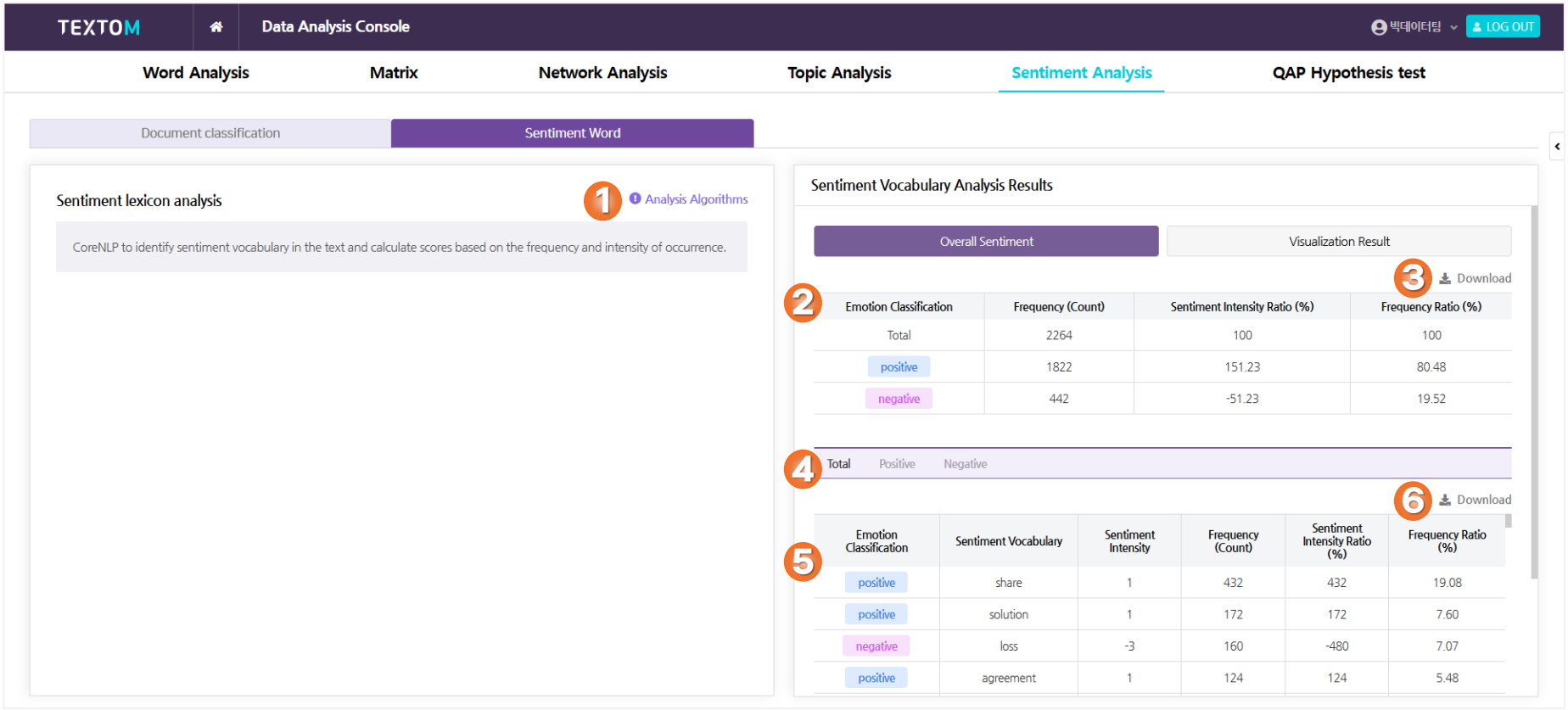


① **Visualization Result**
The results of analysis are presented in a pie chart.

② **Visualization Setting**
Users can modify the attributes of the visualization results as desired.

③ **Zoom In**
Click [View] to display an enlarged version of the selected visualization in a pop-up window.

④ **Download Visualization**
Download the visualization result as a PNG file.



① Selected Algorithm Description
Click the button to view the description of the algorithm.

② Summary of Overall Sentiment Results
View the frequency and percentage results.

③ Download the comprehensive sentiment analysis results as an Excel (.xlsx) file.

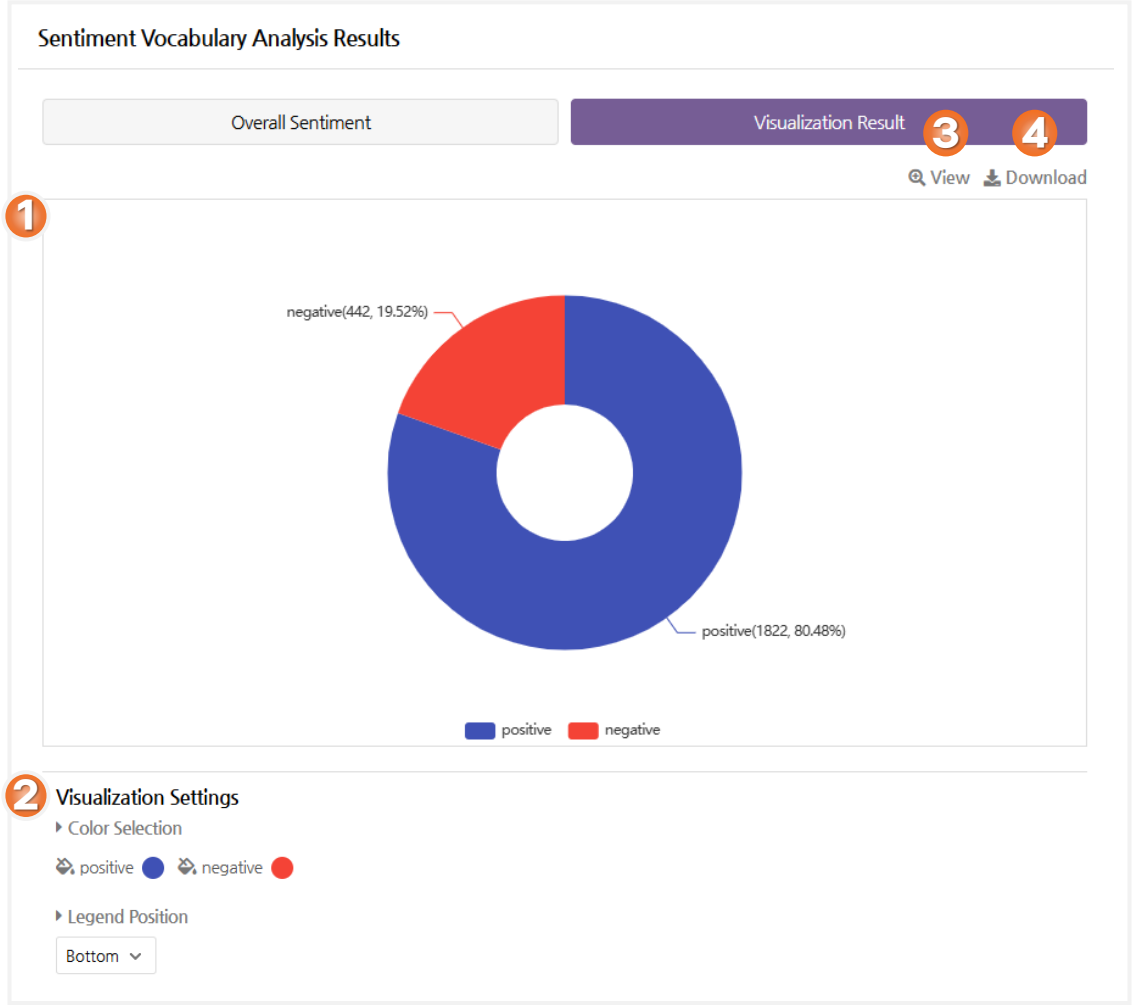
④ Detailed Sentiment Tab
Click the tab for each detailed sentiment to view the sentiment vocabulary.

⑤ View Sentiment Classification Results
View the vocabulary, sentiment intensity, frequency, and percentage for each detailed sentiment.

⑥ Download the detailed sentiment vocabulary analysis result as an Excel (.xlsx) file.

Sentiment Analysis

Displays the visualization results of Sentiment Analysis.

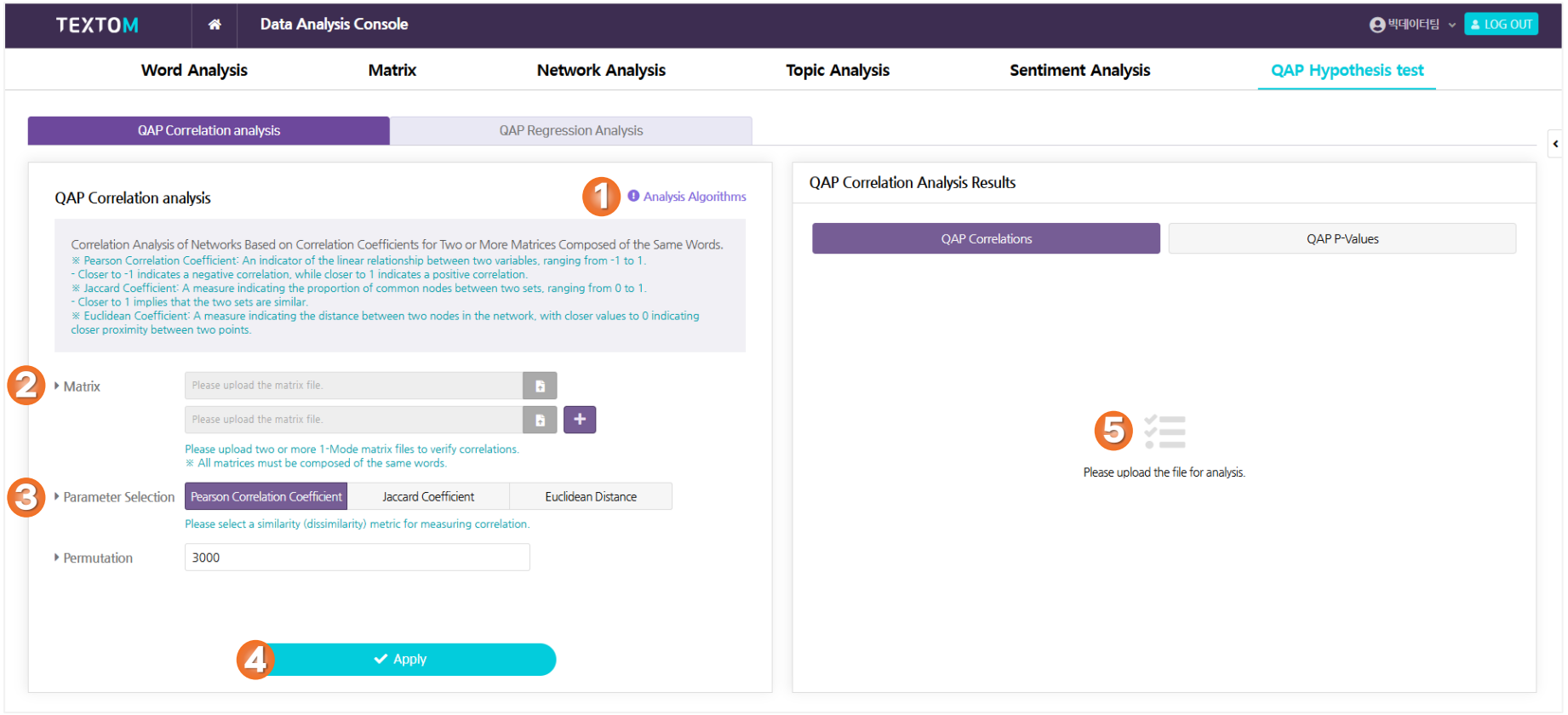


① **Visualization Result**
The results of analysis are presented in a pie chart.

② **Visualization Setting**
Users can modify the attributes of the visualization results as desired.

③ **Zoom In**
Click [View] to display an enlarged version of the selected visualization in a pop-up window.

④ **Download Visualization**
Download the visualization result as a PNG file.



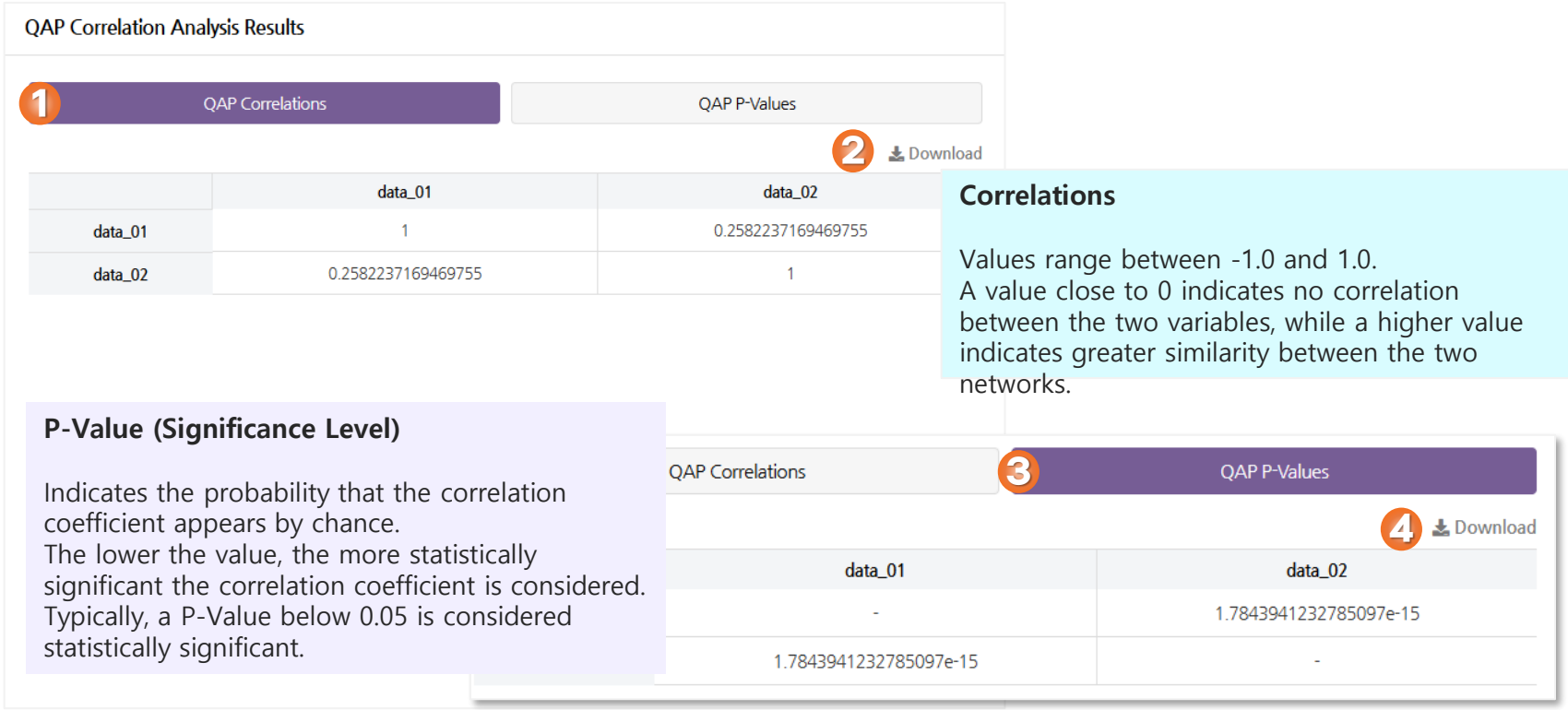
① **Selected Algorithm Description**
Click the button to view the description of the algorithm.

② **Upload the Matrix File**
Upload the matrix for QAP analysis. To check correlations, please upload at least two matrices.

③ **Select Parameters**
Choose [Parameter Selection] and [Permutation] from the parameters.
Tip. Permutation involves randomly rearranging the nodes of the matrix and repeating this process multiple times. It is recommended to perform this at least 1,000 times to evaluate statistical significance.

④ Proceed with the analysis using the uploaded files and selected conditions.

⑤ Display the analysis results.



Correlations

Values range between -1.0 and 1.0.
A value close to 0 indicates no correlation between the two variables, while a higher value indicates greater similarity between the two networks.

P-Value (Significance Level)

Indicates the probability that the correlation coefficient appears by chance.
The lower the value, the more statistically significant the correlation coefficient is considered.
Typically, a P-Value below 0.05 is considered statistically significant.

To understand the similarities and differences between the Korean and overseas markets' preferences for Korean dramas, we conducted a QAP correlation analysis.

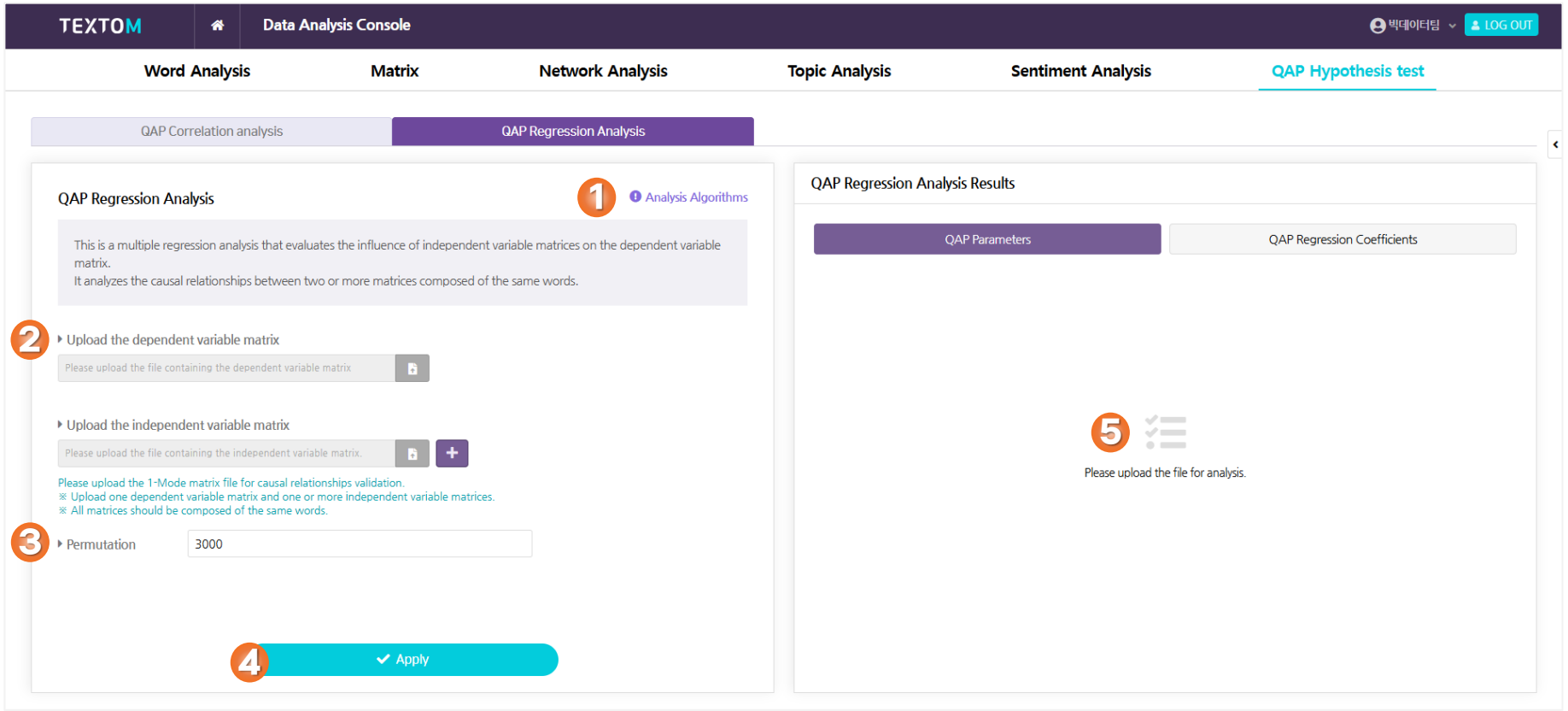
With a correlation coefficient of 0.2582, there is a weak positive correlation between the two networks.
The extremely small P-Value of 2.927564572264821e-17 indicates that the likelihood of this correlation occurring by chance is nearly zero.

① QAP Correlation Analysis Results
Check the QAP correlation analysis results.
Tip. Correlation coefficients only indicate linear relationships. If the relationship between two variables is non-linear, the correlation coefficient may appear low.

② Download Correlation Analysis Results
Download the analysis results as an Excel (.xlsx) file.

③ QAP P-Value Analysis Results
Choose [Parameter Selection] and [Permutation] from the parameters.
Tip. A high correlation coefficient and a low P-Value indicate a strong and statistically significant relationship between the two networks.

④ Download P-Value Analysis Results
Download the analysis results as an Excel (.xlsx) file.



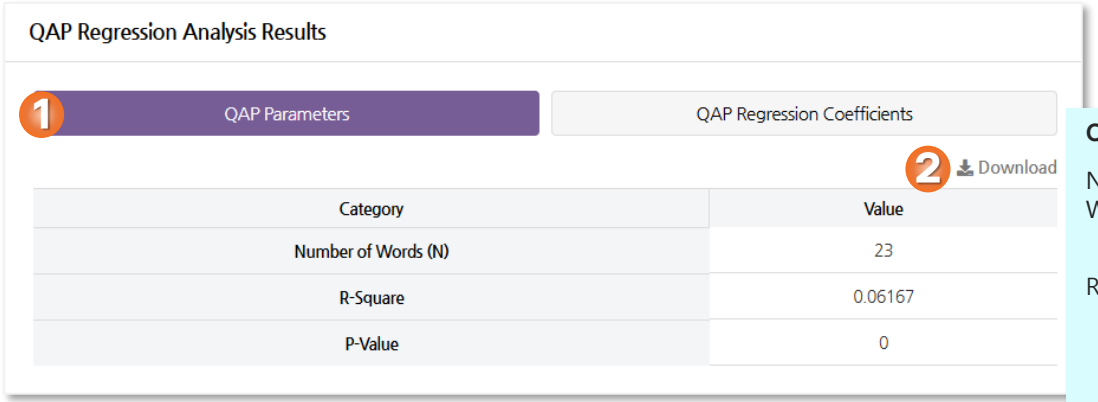
① **Selected Algorithm Description**
Click the button to view the description of the algorithm.

② **Upload the Matrix File**
Upload the matrices for QAP analysis. Please upload one dependent variable matrix and one or more independent variable matrices to check causal relationships.
Tip: The matrices must be 1-Mode matrices, and all matrices should be composed of the same words.

③ **Select Parameters**
Choose [Parameter Selection] and [Permutation] from the parameters.
Tip. Permutation involves randomly rearranging the nodes of the matrix and repeating this process multiple times. It is recommended to perform this at least 1,000 times to evaluate statistical significance.

④ Proceed with the analysis using the uploaded files and selected conditions.

⑤ Display the analysis results.

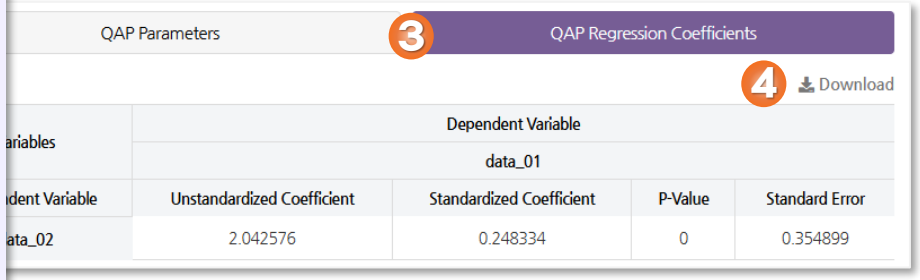


QAP Parameters

Number of Words (N)	Number of variables
R-Square	Indicates the coefficient of determination; the closer it is to 1, the better the model explains the data
P-Value	Indicates the statistical significance of the model; lower values mean that the independent variables of the model have a significant effect on the dependent variable

QAP Regression Coefficients

Unstandardized Cofficients	The effect of independent variables on the dependent variable
Standardized Coefficients	Comparison of the impact between independent variables
P-Value	Statistical significance of the model
Standard Error	Variability and reliability of the regression coefficients

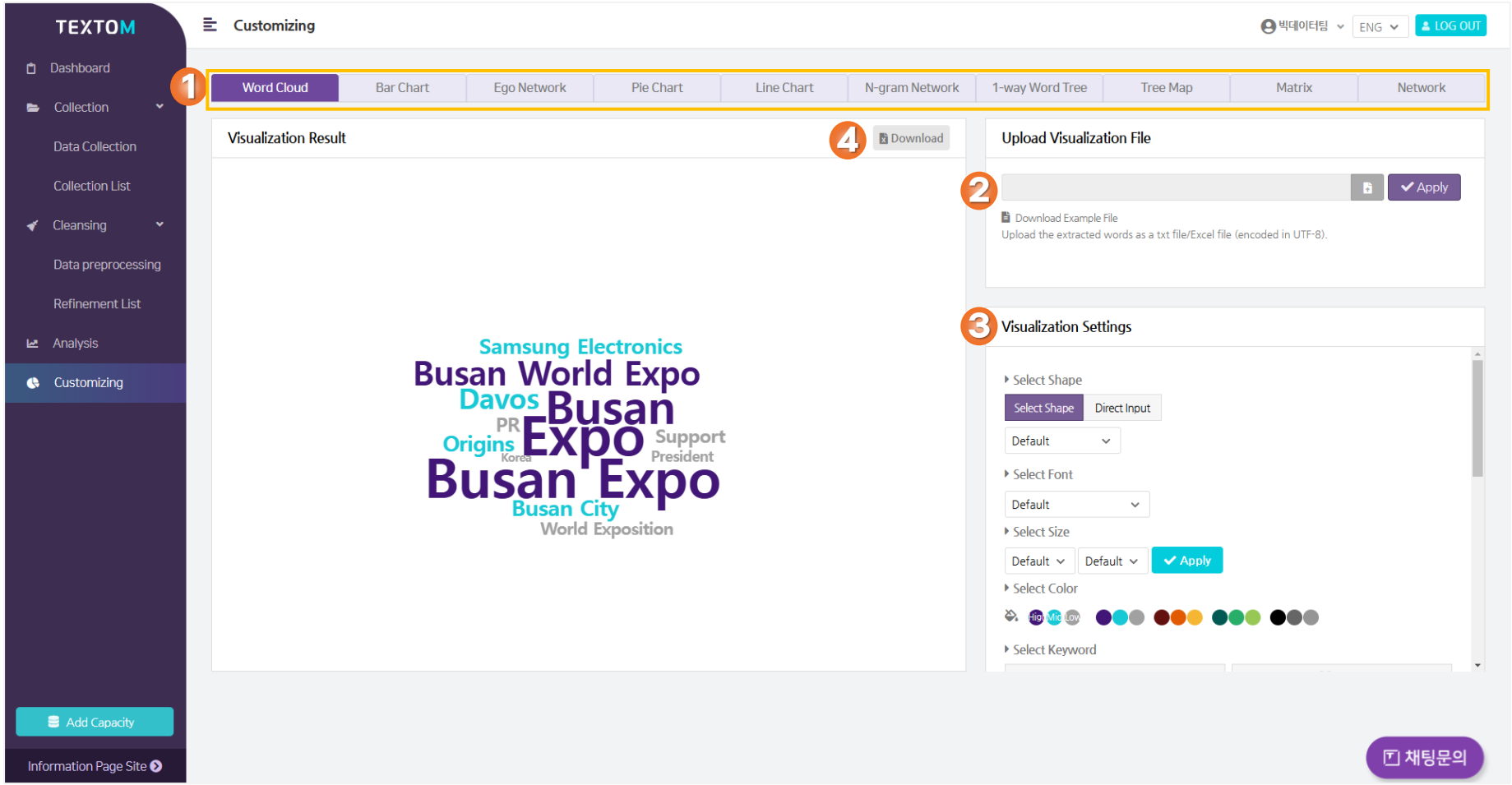


① QAP Parameters Analysis Results
Check the results of the input values analyzed by QAP regression analysis.

② Download QAP Parameters Analysis Results
Download the analysis results as an Excel (.xlsx) file.

③ QAP Regression Coefficients Results
Check the results of the QAP regression coefficients analysis.

④ Download QAP Regression Coefficients Analysis Results
Download the analysis results as an Excel (.xlsx) file.



① **Customizing Type**
Select the type of visualization the user wants to customize.

Word Cloud	N-gram Network
Bar Chart	1-way Word Tree
Ego Network	Tree Map
Pie Chart	Matrix
Line Chart	Network

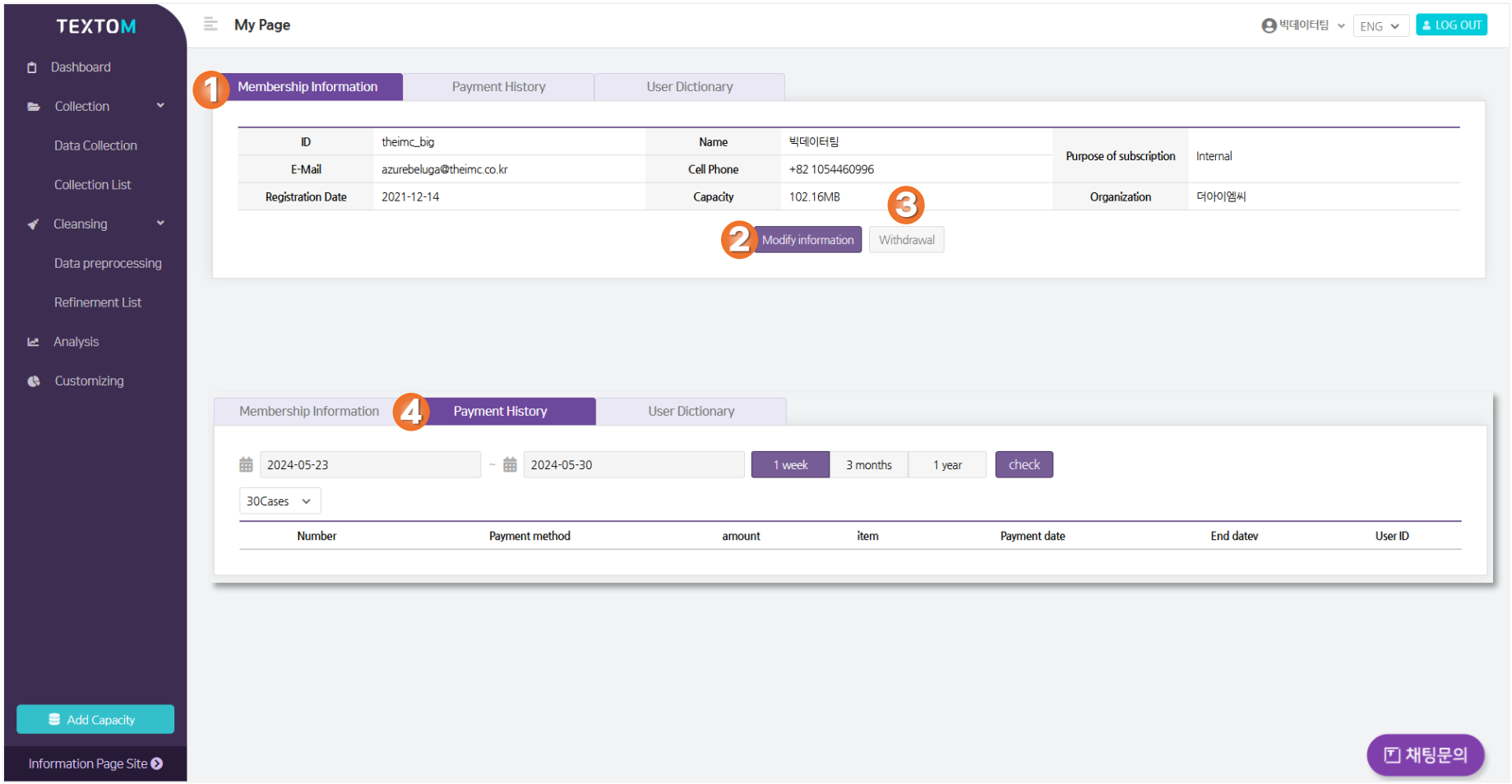
② **Upload Visualization File**
Upload the data file suitable for the chosen type and click the [Apply] button.
Tip. Use [Download Example File] to check the data formatting method.

③ **Visualization Setting**
Users can modify the attributes of the visualization results as desired.

④ **Download Visualization**
Download the visualization result as a PNG file.

My Page

Check your membership information and payment history.

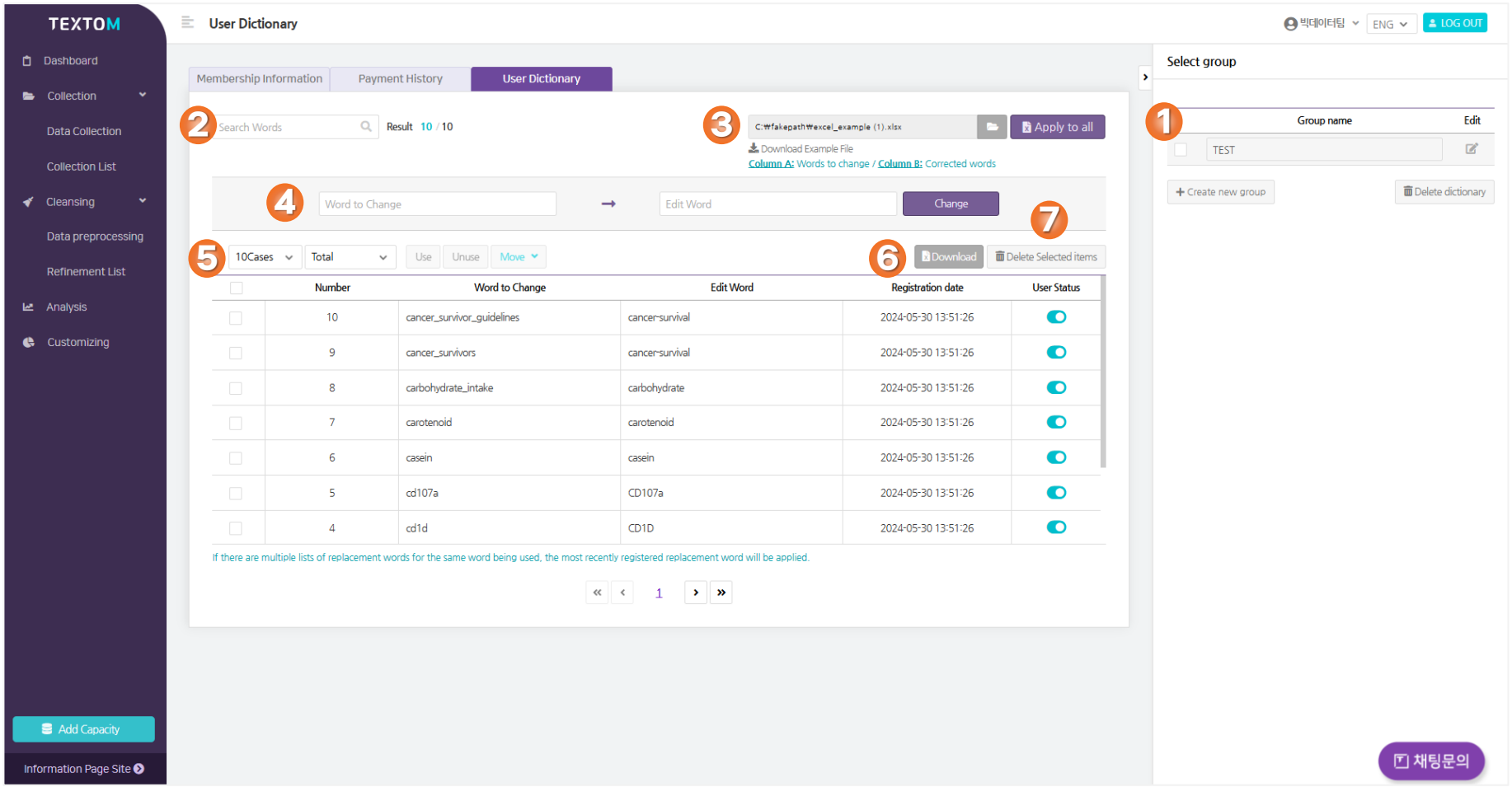


① View Member Information.

② Edit Information
Edit your member information after entering your password.

③ Membership Cancellation
Cancel your membership by entering your ID and password.

④ Payment History
Set a period to view your payment history and related information.



① **Group Selection and Management**
Select a group to manage the user dictionary. Click the [Edit] button to change the group name.

② **Search Words**
Search for words to view the list of modified words that include the searched word.
Tip: Use this to check for duplicate words.

③ **File Upload**
Upload a user dictionary file that you have created and click the [Apply All] button. Use the [Download Sample File] to check the formatting method.

④ **Enter Words**
Input the words to be changed and the modified words.

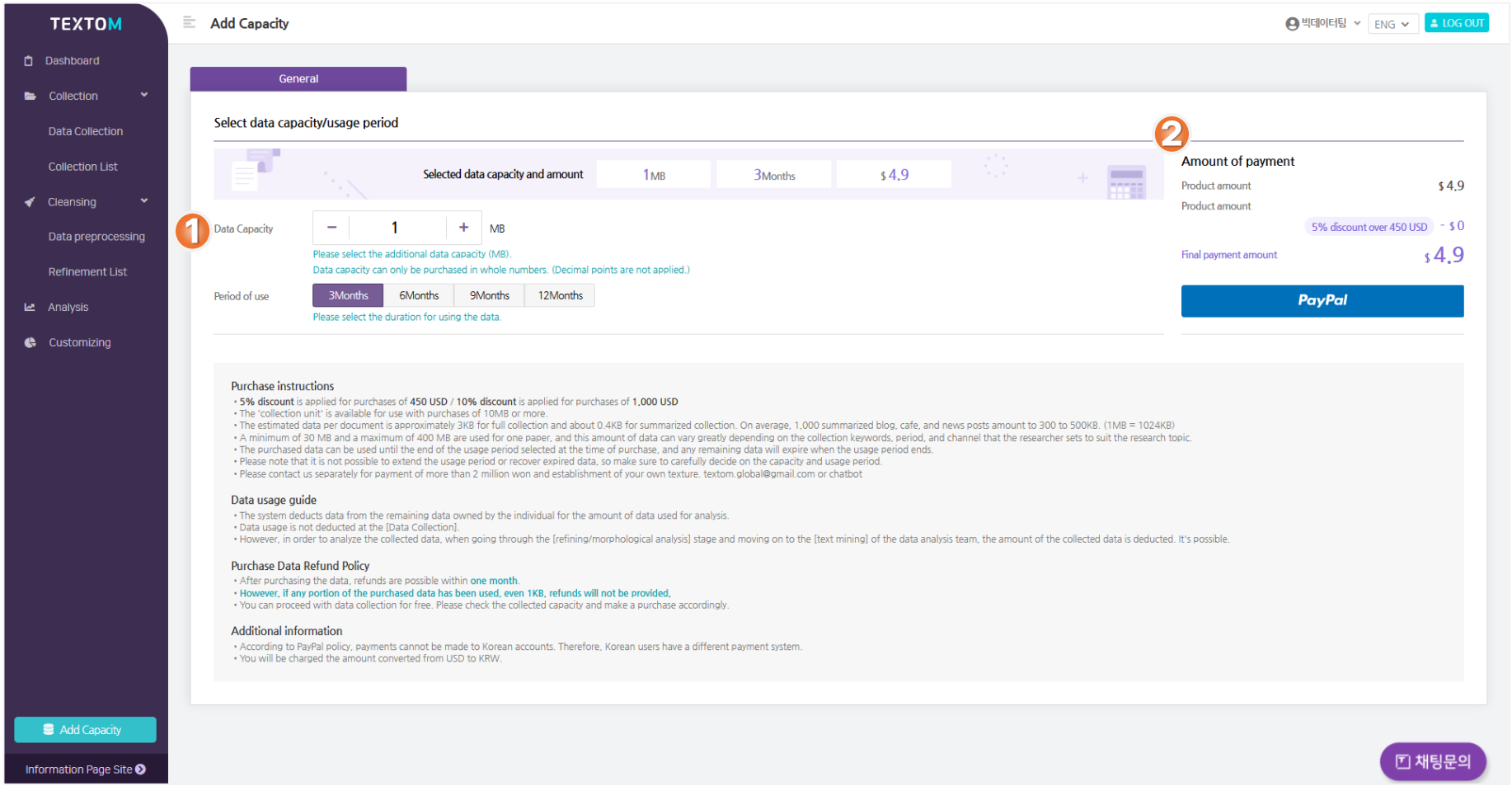
⑤ **Select List View and Usage Options**
Choose the method for viewing the word list and whether to use the words.

⑥ **Download User Dictionary**
Download the user dictionary as an Excel (.xlsx) file.

⑦ **Delete Modified Words**
Select the checkboxes to delete modified words.

Add Capacity

Add the data capacity needed for analysis.



① Select Data Capacity And Usage Period

Choose the additional data capacity and usage period.

TIP. Capacity can be purchased in 1MB increments, and the usage period can be selected in 3-month, 6-month, 9-month, or 12-month intervals.

② Estimated Payment Amount

The estimated payment amount for the selected purchase options is calculated.

TIP. For payments over 1,500 USD and setting up your own TEXTOM, please contact us.