

Setup Kafka clusters in UI

This is a guide to setting up the Kafka clusters using input provided in the text mining UI sidebar.

1. Navigate to the link for the text mining app ex:<http://localhost:7860/>

2. Click on "Choose server" in sidebar

The screenshot shows the Crawler UI interface. At the top, there are tabs: Crawler UI (which is active), Relevance UI, Question Answering UI, and DEBUG/LOG Console. Below the tabs, there's a section titled "Start Crawl Instance" with a play button icon. A "Crawl Identifier" field is present with placeholder text "Unique identifier to manage the crawl. Ex: flood_dortmund_20250611". On the left sidebar, under "Mandatory Kafka Setup" (with a warning icon), there are several configuration fields:

- "Choose server": A dropdown menu with "crexdata-unauth" selected, highlighted by an orange border.
- "Bootstrap server": An input field containing "server.crexdata.eu:9192".
- "kafka topic with relevant tweets": An input field containing "text_mining_processed_relevant_tweets_stream".
- "kafka topic sending to ARGOS": An input field containing "Tango".

At the bottom right of the main area, it says "Created with Tango.us".

3. Set to "crexdata-auth" if using authenticated server.

Note: The authentication settings should have been previously configured. See README.md

The screenshot shows the CREXDATA setup interface. On the left, there's a logo and the text "Critical Action Planning over Extreme-Scale Data". Below it, a section titled "Mandatory Kafka Setup" contains a warning icon. A note says "Enter kafka server details and click connect kafka". A dropdown menu labeled "Choose server" has two options: "crexdata-unauth" (selected) and "crexdata-auth". The "crexdata-auth" option is highlighted with a red box. Below the dropdown are two input fields: "kafka topic with relevant tweets" containing "text_mining_processed_relevant_tweets_stream" and "kafka topic sending to ARGOS" containing "text_mining_to_argos_stream". On the right side, there are three input fields with placeholder text: "rtmund_20250611", "X password for crawling account password" (containing a redacted password), and "Gmail app access password" (containing a redacted password). A note below the first field says "Should be the app pass for the crawl". On the far right, a "Crawl language" field is set to "German". At the bottom left is a "Tango" watermark, and at the bottom right is the text "Created with Tango.us".

4. Select "crexdata-unauth" if using an unauthenticated server.

This screenshot is identical to the one above, showing the CREXDATA setup interface. The "Choose server" dropdown now has "crexdata-auth" selected (highlighted with a blue box) instead of "crexdata-unauth". The other UI elements, including the Kafka topics and the right-hand configuration fields, remain the same.

5. The "Bootstrap server" field becomes visible if using an unauthenticated server.

Fill in the bootstrap server IP and port, ex: localhost:9092

Critical Action Planning over Extreme-Scale Data	rtmund_20250611	X password for crawling account password
Mandatory Kafka Setup ⚠		
Enter kafka server details and click connect kafka		
Choose server crexdata-unauth	Bootstrap server server.crexdata.eu:9192	Gmail app access password Should be the app pass for th.....
kafka topic with relevant tweets text_mining_processed_relevant_tweets_stream	r is of a running crawl	Crawl language Only those supported by X German
kafka topic sending to ARGOS text_mining_to_argos_stream		
kafka topic with relevant tweets for QA text_mining_processed_relevant_tweets_stream		
Tango	Created with Tango.us	

6. In the "Kafka topic with relevant tweets" field, you can set the Kafka topic used in the "Relevance UI" tab to get posts relevant to a weather crisis.

Note: This topic needs to already have posts classified as "fire" or "flood" using the event type prediction model.

Mandatory Kafka Setup ⚠		
Enter kafka server details and click connect kafka		
Choose server crexdata-unauth	Bootstrap server server.crexdata.eu:9192	Gmail app access password Should be the app pass for th.....
kafka topic with relevant tweets text_mining_processed_relevant_tweets_stream	r is of a running crawl	Crawl language Only those supported by X German
kafka topic sending to ARGOS text_mining_to_argos_stream		
kafka topic with relevant tweets for QA text_mining_processed_relevant_tweets_stream		
kafka topic crawler commands		
Tango	Created with Tango.us	

7. In the "Kafka topic sending to ARGOS" field, you can set the Kafka topic to send

posts to ARGOS.

Note: This topic should be monitored by the ARGOS app

The screenshot shows a configuration interface for a Kafka-based system. On the left, there's a dropdown menu set to "crexdata-unauth". Below it, a "Bootstrap server" field contains "server.crexdata.eu:9192". A section for "kafka topic with relevant tweets" lists "text_mining_processed_relevant_tweets_stream". Another section for "kafka topic sending to ARGOS" lists "text_mining_to_argos_stream", which is highlighted with a red box. A third section for "kafka topic with relevant tweets for QA" lists "text_mining_processed_relevant_tweets_stream". Below these are sections for "kafka topic crawler commands" ("text_mining_crawler_commands") and "kafka topic with crawler logs" ("text_mining_crawler_logs"). On the right side, there's a "Crawl language" section with "German" selected. At the bottom, the "Tango" logo is on the left and "Created with Tango.us" is on the right.

8. In the "Kafka topic with relevant tweets for QA" field, you can set the topic used in the "Question Answering UI" tab to get posts relevant to a weather crisis for question answering.

This screenshot is similar to the previous one but focuses on the "kafka topic with relevant tweets for QA" field. The "text_mining_processed_relevant_tweets_stream" entry is highlighted with a red box. The rest of the interface, including the bootstrap server, other Kafka topics, and crawl language settings, remains the same.

9. In the "Kafka topic crawler commands" field, you can set the topic to send crawler commands.

Note: This topic should be configured in crawler component. (see X_Crawler)

kafka topic sending to ARGOS
text_mining_to_argos_stream

kafka topic with relevant tweets for QA
text_mining_processed_relevant_tweets_stream

kafka topic crawler commands
text_mining_crawler_commands

kafka topic with crawler logs
text_mining_crawler_logs

kafka topic with processed tweets for latency
text_mining_processed_tweets_stream

Show latency plot

Tango

Created with Tango.us

10. In the "Kafka topic with crawler logs" field, you can set the topic to get crawler logs.

Note: This topic should be configured in crawler component. (see X_Crawler)

text_mining_to_argos_stream

kafka topic with relevant tweets for QA
text_mining_processed_relevant_tweets_stream

kafka topic crawler commands
text_mining_crawler_commands

kafka topic with crawler logs
text_mining_crawler_logs

kafka topic with processed tweets for latency
text_mining_processed_tweets_stream

Show latency plot

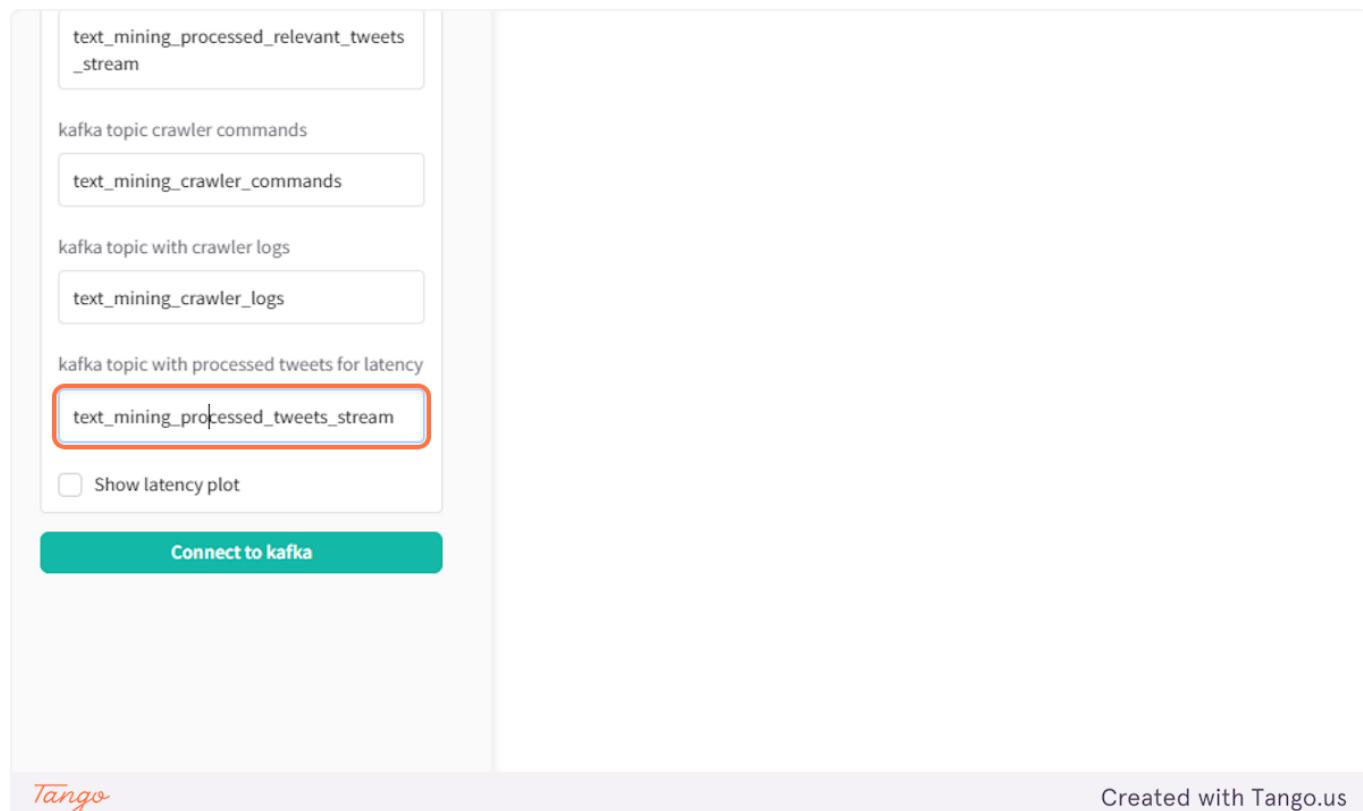
Connect to kafka

Tango

Created with Tango.us

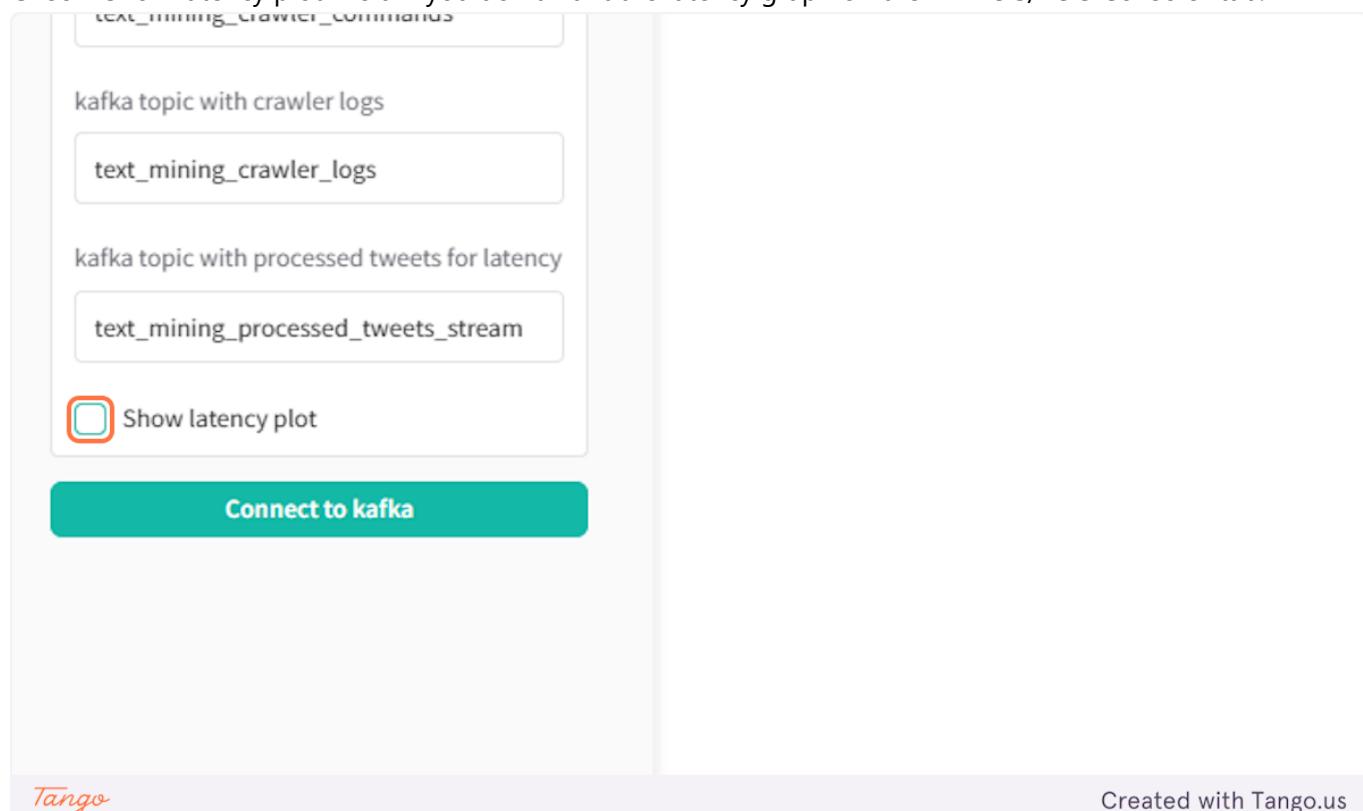
11. In the "kafka topic with processed tweets for latency" field, you can set the topic with post processed by the event type prediction component. These posts have latency calculation and

allow viewing a latency graph, for debug purposes.



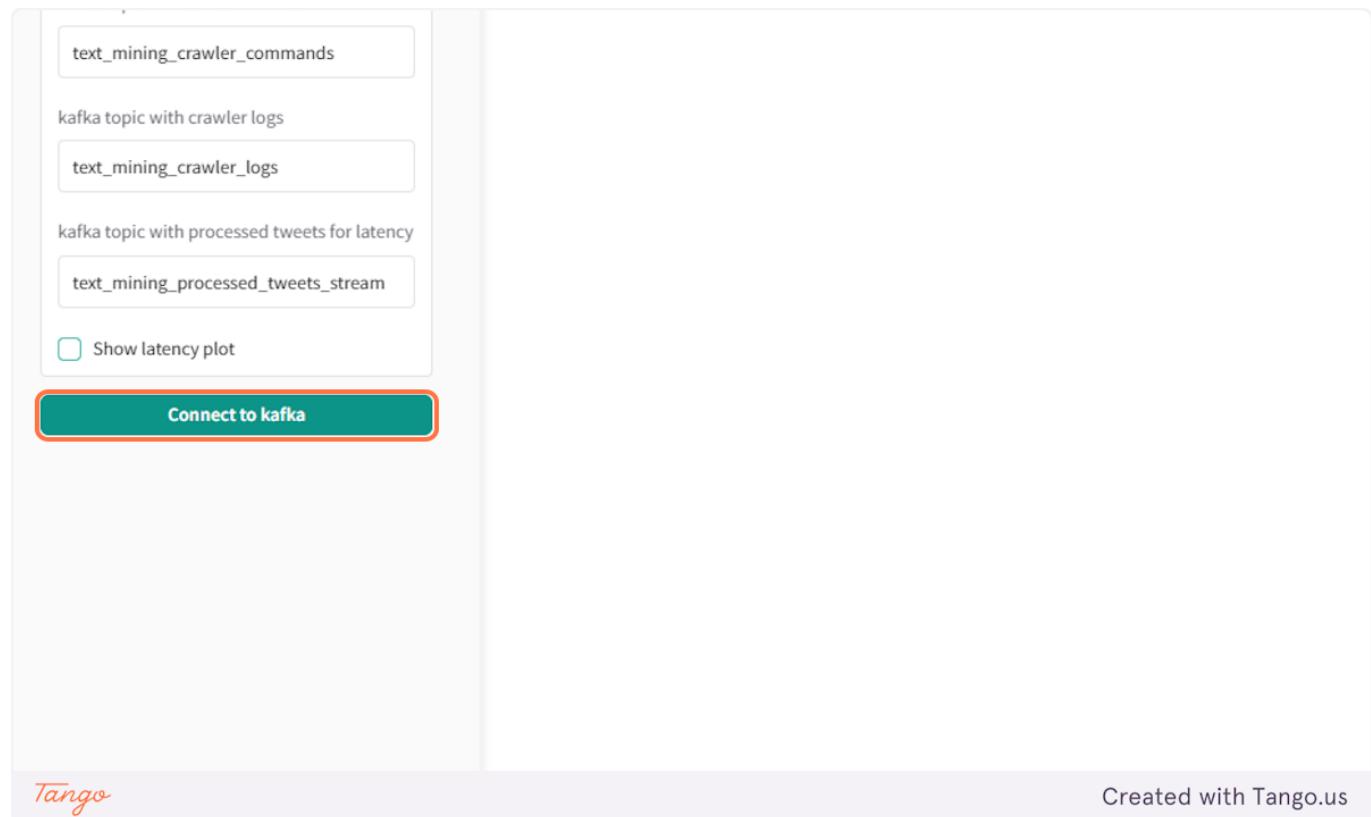
12. Uncheck "Show latency plot" field if you don't want the latency graph on the "DEBUG/LOG Console" tab.

Check "Show latency plot" field if you don't want the latency graph on the "DEBUG/LOG Console" tab.



13. Click "Connect to kafka" to finalize the config process and connect to kafka

clusters.



Created with [Tango.ai](#)

Created with Tango.us