Overview

TGForge is a Streamlit-based application that collects data from one or more public Telegram channels (or groups) using Telegram's official API. With TGForge you can extract channel details, messages, forwards, and participant information for further analysis.

- Tool here: https://isd-tgforge.streamlit.app/

Prerequisites

Telegram API Credentials:

- Obtain your API ID and API Hash from Telegram's official page - https://core.telegram.org/api/obtaining_api_id
- For guidance, watch this overview video in how to access your API ID and API hash - https://www.youtube.com/watch?v=tzYTLjdr7rI

Account Considerations:

- You may use a burner account; however, note that new accounts sometimes cannot obtain API credentials.

Step 1: Enter Telegram API Credentials

Input Credentials:

- Enter your API ID, API Hash, and phone number in the provided fields.
- Click the Next button.

Issues:

- If you encounter errors (e.g., "database is locked"), click Refresh Session and try again. If issues persist, contact the DAU.
- **Note:** A successful submission will automatically move you to Step 2, after a few seconds.

Step 2: Authentication

Check Your Telegram App:

- Open Telegram on the account associated with your API credentials. You should receive a login code as if you were signing in on a new device.

Enter the Code:

- Input the received verification code in the app.
- A notification may appear on your phone indicating an attempted login. Rest assured, this is only for authentication purposes and cannot be used to access your account.

Authenticate

- **Note:** Once authenticated, the app will automatically advance to Step 3.

Step 3: Using TGForge

1) **Channel Info**

a. **What It Does:** Retrieves basic channel details such as alternative names, type (group/channel), and creation date.
2) **Messages**
    a. **What It Does:** Collects all messages from the selected channel(s) or group(s).
    b. **How to Use:** Separate multiple channels with commas (e.g., durov, washingtonpost). By default, it collects all posts. You can optionally filter by a specific date range.
    c. **Output:** Download options are available for both a CSV file (raw messages) and an Excel file (analytics).
3) **Forwards**
    a. **What It Does:** Similar to message collection but focuses on forwarded messages only.
4) **Participants**
    a. **What It Does:** Retrieves group/channel members, either directly via the API or by extracting senders from messages.
    b. **Two Methods:**
        i. **Default:** Pulls participants directly from the API.
        ii. **Via Messages:** Collects participants based on message activity within an optional date range, supplementing API data.
    c. **Note:** Large or highly active groups might take longer to process. For extensive data pulls, consider scanning groups one at a time or contact the DAU.

Running a Scan

Initiate Scan:
- After selecting your scan type (Channel Info, Messages, Forwards, or Participants) and entering channel names, click the respective fetch button.

Interrupting a Scan:
- Press 'Refresh / Cancel' to stop an ongoing data pull.

Additional Notes

- **Processing Time:** TGForge is efficient but may take significant time for large data sets.
    o Make sure your computer stays awake and connected to the internet. Loss of internet, going to sleepmode, etc. will interrupt a download and you will need to start over. Make sure to save CSVs/XLSX files if desired, as similarly even once a scan has been completed you may similarly lose your data.
    o For large-scale collection, contact the DAU.
- **Security:** The API credentials you enter are solely for data extraction. They cannot be used to access your account.

- **Privacy:** Channels or groups do not receive any indication that they have been scanned.
- **Limitations:** TGForge currently works only on public channels.
- **Support:** For bugs or feature requests, contact Nathan or a member of the DAU.