DEMOCRACY REPORTING INTERNATIONAL

METHODOLOGY

The research method used combines quantitative data analysis methods with qualitative content analysis. Quantitative data collected on a monthly basis is aggregated from Facebook's Graph Application Programming Interface (API) through three third-party tools: Netvizz, Buzzsumo and Crowdtangle.

Queries are built based on known keywords associated with predefined topics. These queries are updated on a weekly basis depending on relevance. URLs of relevant news articles are retrieved through Buzzsumo's Content Analyser tool by searching for queries in thematic categories related to the political transition in Libya and to DRI's mandate. These thematic categories include: (1) the constitution, (2) elections, (3) the UN roadmap, and (4) security and legitimacy. The engagement metrics extracted from Buzzsumo's Content Analyser include the total number of Facebook engagements.

To receive a more detailed breakdown of engagement metrics, URLs that are retrieved from Buzzsumo based on the keyword queries, are inserted into Netvizz' Link Stats tool that provides the number of likes, reactions, shares and comments at the time of extraction.
The Facebook-owned Crowdtangle browser extension is sometimes used on specific URLs to analyse which Facebook pages that were sharing the content the most.

To qualitatively analyse conversations and comments, Netvizz is used to extract the top 200 comments of the selected Facebook posts. Then 20% of these comments are manually selected based on a random selection method. Redundant comments that only have one word replies or are not relevant to the post are removed.

The remaining comments are then coded according to a list of predefined codes. The coding dictionary was created with the following categories: 1) thematic codes capturing the main topics covered, 2) narrative codes capturing sentiments in support of or opposition to certain beliefs, attitudes, or groups 3) emotive codes capturing positive, negative and neutral tones 4) language codes capturing abuse or religious sentiments. The coding dictionary is also downloadable on the report page.

The Facebook search tool is also used to manually collect a list of public Facebook posts relating to the four thematic categories of interest. From the manual Facebook search results, Facebook engagement numbers in each relevant post are collected. The engagement metrics of these posts are entered in a spreadsheet.

For Twitter data collection, the Twitter Archiving Google Sheet (TAGS Hawksey 2014) is used. This sheet (with built-in scripts) automatically collects tweets originating from a specific hashtag and archives them with other metadata including username, timestamp, followers count, and location.

Restrictions by the Twitter Archiving Google Sheet limit data collection to 7 days after a tweet is posted. Selected hashtags are extracted as soon as they become popular to overcome this limitation.

Limitations

Data collected via Netvizz and Buzzsumo is public data and does not include clicks or reach or impressions. The data pulled by the social media tools employed by DRI is not disaggregated by gender, age or location. This is because Facebook has banned third parties from accessing users' demographics data to protect user information on the platform.

The Facebook-specific gender related information in this report is gathered and analyzed manually. Due to the virtual absence of women commenting on public pages, Twitter has been used instead given its popularity as a safe space for women to interact online. The report ensures that the content analysis is based on public pages and profile pages. It does not by any means divulge or expose information based on posts in private fora or private user profiles without their prior consent.

The qualitative sampling method outlined above represents another limitation. Given the large sample, a selection (of a randomized 20% of the 200 top comments) must be made to reduce the data that is manually coded. In consequence, some trends in comments which were not selected for coding might be overlooked.