

# *Transforming Crisis Narratives on Social Media into Open Knowledge*

---

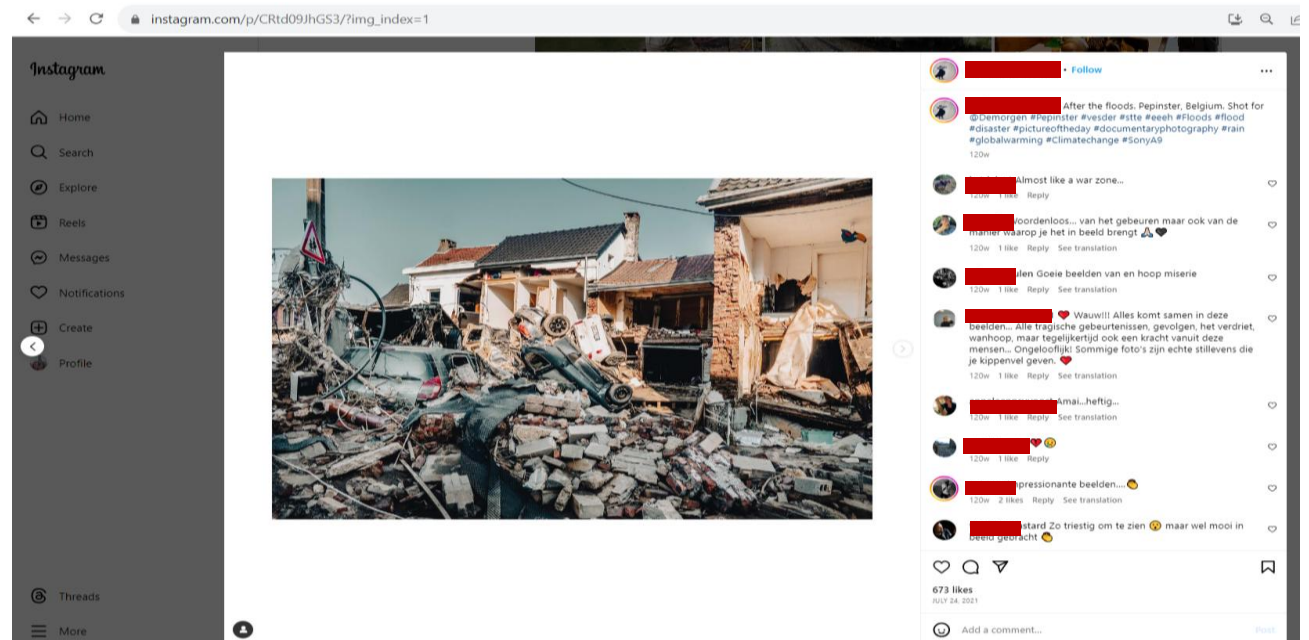
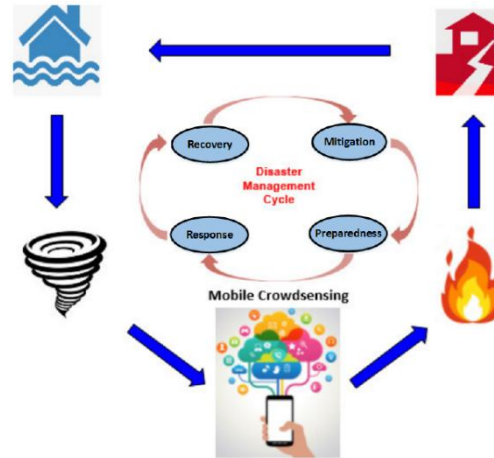
LEMA

Manal Ginzarly, Ph.D.  
Marie Skłodowska-Curie Actions (MSCA)  
27 Nov 2025

# From Public Posts to Open Data

User-generated content,  
Publicly available data posted in real time (within minutes of an event) on platforms like Twitter/X, Facebook, Instagram, and TikTok

Posts come with **rich metadata**:  
geographic locations, timestamps, text, photos, videos, and engagement metrics such as likes and comments.



# Data Acquisition and pre-processing

# Open and Transparent Data Acquisition

Using **APIs** (Application Programming Interfaces) and **data-scraping tools**, social media metadata can be collected at scale.

By publishing the scraper code and workflow openly (e.g., on GitHub), the entire data acquisition process becomes transparent.

Ensuring compliance with the **FAIR** (**F**indable, **A**ccessible, **I**nteroperable, **R**eusable) principles for future reuse and reproducibility by the scientific and professional community.

## GitHub Example

The screenshot displays a GitHub repository interface. At the top, navigation links for 'Projects', 'Security', and 'Insights' are visible. Below this, the repository name 'Ginzarly' is shown with a branch selector set to 'main', '1 Branch', and '0 Tags'. A search bar and a 'Code' button are also present. The commit history table lists the following files and their commit details:

File	Commit Message	Commit Date
LICENSE	Initial commit	4 months ago
README.md	Update README.md	4 months ago
instagram_scraper.py.txt	Add files via upload	4 months ago
sentiment_analysis.py	Add files via upload	4 months ago

The README section is titled 'LANDSCAPEforCHANGE' and features the following content:

### instagram\_scraper.py

#### Features

- Scrapes Instagram posts for a specific location (e.g. Verviers) or hashtag (e.g. #Belgiumflood)
- Downloads images from the posts
- Extracts captions, likes, comments, hashtags, user info, and timestamps
- Saves all data into structured CSV and JSON formats
- Avoids duplicate posts and images
- Scrapes up to 50 comments per post using the Instagram Comment Scraper

#### Requirements

- Python 3.x
- An [Apify](#) account and API token
- The following Python libraries:
  - requests
  - csv
  - os
  - apify-client
  - json

# Data Pre-processing

## Pre-processing for open, reusable social media metadata

**Standardizing formats:** harmonizing timestamps, converting coordinates, and unifying language/encoding.

**Structuring data:** organizing text, images, and metadata into open formats (ex CSV) with clear, well-defined schemas.

**Anonymization & ethical filtering:** removing personal identifiers to ensure GDPR-compliant open data.

Creating README files and metadata descriptions to support transparency, reproducibility, and reuse.

## Mendeley Example

### Social Media Use in Disaster Response: Empowering Communities Resilience









Published: 29 July 2025 | Version 2 | DOI: 10.17632/78h67jgfrt.2  
Contributor: Manal Ginzarly

#### Description

Social media metadata is employed to investigate the use of two social media platforms, Facebook and X, during the 2021 flood in Belgium, both in the immediate aftermath of the crisis and in the short-term. First, Facebook data is employed to analyze the activities of Facebook community groups established post-crisis over a six-month period. Then, the tweets related to the 2021 flood in Belgium are analyzed. The analytical framework employs (1) a social media data-driven quantitative and qualitative text analysis using topic modeling and sentiment analysis. The findings highlight the different roles played by the two social media platforms. Facebook served as an effective platform to mobilize and organize local communities for immediate and practical support, while Twitter served as a platform for broader global engagement and advocacy. The convergence of results from diverse data sources provides comprehensive insights into the effectiveness and challenges of leveraging social media for community resilience in the aftermath of disaster events.

[Download All 912 KB](#) ⓘ

#### Files

 Belgium Flood Tweets user profile.xlsx	81.9 KB	<a href="#">↓</a>
 Entraide inondations Li_ge.xlsx	327 KB	<a href="#">↓</a>
 Les petites fourmis de la Hoegne.xlsx	145 KB	<a href="#">↓</a>
 Particules solidaires de la vesdre.xlsx	40.4 KB	<a href="#">↓</a>
 People against flows.xlsx	136 KB	<a href="#">↓</a>
 River Cleanup.xlsx	35.8 KB	<a href="#">↓</a>
 stats fb members.xlsx	11 KB	<a href="#">↓</a>
 tweets belgium flood raw data.xlsx	154 KB	<a href="#">↓</a>

Usage	
Downloads:	117
Views:	34

 [View details >](#)

#### Latest version

Version 2  
Published: 29 Jul 2025  
DOI: 10.17632/78h67jgfrt.2

#### Cite this dataset

Ginzarly, Manal (2025), "Social Media Use in Disaster Response: Empowering Communities Resilience", Mendeley Data, V2, doi: 10.17632/78h67jgfrt.2

[Copy to clipboard](#)

#### Previous versions

Version 1 • 9 June 2025

#### Version comparison

[Compare versions](#)

# Ethical Considerations

**Respect for user privacy:** no attempts to deanonymize users or track individuals.

**Minimization of sensitive information:** remove names, faces, or exact coordinates when not essential to analysis.

**Clear documentation:** provide disclaimers, data processing logs, and limitations.

**No redistribution of raw private media:** instead, share metadata or blurred/processed versions.

Comply with **GDPR**, **platform terms**, and **MSCA ethics requirements**.