LINKS

Strengthening links between technologies and society for European disaster resilience

White Paper
Legal Assessment of Data Harvesting - use of Social Media Crowdsourcing in Europe

Case of Denmark

Authors:
Badar Shah
Zaynab Naji
University College Copenhagen

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 883490

Introduction

The purpose of this document is to provide a legal insight into the data harvesting of information from social media, during hazards, emergency incidents and crises situations. The legal insight will focus on both national and international rules, when and how they apply. Furthermore, there will be a focus on specific social media platforms and their individual terms of service - some provide opportunities while others present hindrances.

This document can thus be carved into three sections (in corresponding order) – The national rules that apply in Denmark, the international rules in the form of the General Data Protection Regulation (GDPR) that applies throughout Europe and finally the individual social media platforms themselves. The first section is primarily representative of the legal state in Denmark, while the second section can be used throughout Europe and finally the third section is applicable to any country.

For the sake of understanding, the following definitions are used: Data harvesting or data crowdsourcing is the process whereby a party collects data - either manually or by an automated process - from a platform where the data is available. The collected data can be used for a number of purposes. As the focus of this document is data crowdsourcing from social media, the term "social media crowdsourcing", or SMCS for short, is used.

The term data harvesting should be understood as a supercategory, while SMCS should be understood as a subcategory or a specific way of data harvesting.

The legal basis – Denmark

In addition to the international rules and obligations, Denmark also has national legislation that regulates the proper and legal way of handling data. These rules apply to both public authorities and to private actors (in certain cases). The following explains which legal rules are relevant in which cases.

Scope of administrative law

Danish administrative law sets out a number of duties and rights for authorities and citizens, respectively. Among the most important of these duties for the public authorities can be considered the duty to take notes and to keep records, which are specified in the Danish Public Administration Act. It is thus a duty for any public authority to document and record (save) the information and documents that have formed the basis for the specific decisions carried out by the authorities. The information and the documents must also be named and sorted in a clear manner so that they can be easily retrieved, for example in the event of a request for access to these documents.

It is also important to note that it is generally not allowed to delete logged material that has been saved by the public authorities. This applies regardless of whether the material in question has already fulfilled the purpose for which it was collected. Instead, the material must be archived in accordance with the applicable rules in archive legislation, cf. section 14 of the Danish Data Protection Act. These rules differ

significantly from the general rules on personal data protection, which state that data must be deleted as soon as possible after the purpose of collection has been achieved. These rules are explained in more detail below, in the section on personal data protection.

In addition to the written legal texts mentioned above and the associated preparatory works, applicable administrative law principles can be derived from statements from the Ombudsman, administrative law decisions and, not least, from case law at the civil courts.

Administrative law regulates the relationship between authorities themselves and between authorities and citizens in connection with administrative activities. The activities of public authorities can be divided into several types of actions, but the most relevant ones for this paper can be condensed into decisions, procedural decisions, and actual administrative activities. These can be described as follows:

Decisions
- Most often written, unilateral communications from the authority, directed at either one or more citizens. The decision states what should be the applicable law in a specific situation. For example, this could be a decision that a citizen is entitled to a social benefit.

Procedural decisions
- Decisions made internally by an authority as part of the preparation of a case or the actual organization of how the case should processed are called procedural decisions. Examples include decisions regarding the commencement of case processing or the transfer of a case from one caseworker to another.

Actual administrative activities
- Actual administrative activities consist of the practical execution of public authority tasks. These are not legally binding decisions, but the actual performance of tasks such as patient care, childcare, cleaning, etc.

In practice, it can be difficult to distinguish between the different types of actions because actual administrative activities may result in a final decision, and a decision is dependent on procedural decisions. However, the distinction is important, as a number of rules from the Danish Public Information Act and the Public Administration Act, including the rules on the duty to take notes and the duty to keep records as described above, only apply in connection with decision-making activities. It is expected that data harvesting or SMCS will to a very large extent be considered to fall under the concept of actual administrative activities and not decision-making activities.

Data harvesting and SMCS carried out by private companies, including NGOs, is not covered by administrative law and the rules mentioned above, as the actions only concern private actors and does not include public authorities.

In relation to data harvesting or SMCS carried out by public authorities, this is only covered by administrative law if the action can be considered to be an administrative activity, see above.

It has previously been clarified in the Ombudsman's case law that the use of social media, including Facebook, by public authorities is to be regarded as an administrative activity and that the general rules and principles of administrative law apply in this connection.[1].

[1] Administrationen af Skattestyrelsens Facebook-profil (ombudsmanden.dk)

Public authorities that use SMCS for crisis management, such as the Belgian authorities during the terrorist attack in Brussels in 2016 (see e.g. Marynissen 2020[2]), would thus have to comply with administrative law rules in this regard, if the situation had happened in Denmark.

Lastly, the Danish rules do not differentiate between manual or automatic gathering of information.

Data protection rules - Europe

The EU legislation regarding data protection concerns the protection of personal data. Personal data is defined as data that can either be directly or indirectly linked to a specific person, cf. GDPR art. 4 (1). Personal data can thus range from names and pictures (directly linked) to phone numbers and license plates (indirectly linked).

SMCS without the use of personal data

If all personal data is discarded and only anonymous and/or non-personally identifiable data is used, the data protection rules do not apply. This means that there is no legal obstacle to data harvesting and SMCS as long as no personal data is used.

SMCS with use of personal data

If data harvesting and SMCS will involve the processing of information such as names, images, and addresses, the GDPR and national legislation regarding data protection will apply.

According to GDPR art. 5 (1)(a), all personal data must be processed lawfully, fairly and in a transparent manner in relation to the citizen to whom the data concerns. In relation to legality, it should be noted here that the legal authority and legitimacy for processing this data can be in the form of national law, in the form of the GDPR or in the form of a consent by the citizen.

At present, there are no immediate legal basis in the Danish legislation, nor in EU legislation, which regulates SMCS in relation to crisis management. It is assumed that it will not be possible to obtain individual consent from the individual citizen, since data harvesting must take place automatically across a social media platform, and thus it will not be in accordance with current legislation to carry out SMCS if personal data is also harvested.

SMCS will therefore only be considered legal if any personal data can be sorted out before the data processing takes place or if there is a change in the national legislation and a separate authorization for data processing is drawn up in these specific situations.

In addition to the criterion of legality, the data ethics guidelines must also be considered. These guidelines imply digital responsibility which is about collecting, storing, using, and sharing data in a responsible manner. The consideration behind these guidelines is to create a trusting relationship between the citizen and the data processor. The data processor should not discriminate nor create inappropriate consequences in any other way for the citizenry.

[2] Marynissen, H. & M. Lauder (2020). Stakeholder-Focused Communication Strategy During Crisis: A Case Study Based on the Brussels Terror Attacks. In International Journal of Business Communication, Vol. 57(2) 176–193

Data ethics is not regulated by the GDPR but concerns the good practice in relation to the processing of data. These data ethical considerations are not only put upon public authorities but also private actors.

Principles for processing personal data

When an actor - whether private or public - processes personal data, a number of obligations and requirements must be met. Below is a brief description of the most important obligations and requirements under the GDPR. It should be noted that these only apply if personal data is used. SMCS without the use of personal data is thus not covered by these rules, see the section above.

Storage limitation and data minimization

It follows from Art. 5(1)(c) that only the data that is absolutely necessary to achieve the purpose of the collection may be used. Thus, it is not permitted to collect or process data just for the sake of it or any such unnecessary reason. It further follows from Art. 5 (1) (e) that data must be deleted or anonymized as soon as possible after the purpose of the collection has been achieved, and that personal data may not be stored for longer than strictly necessary.

It should be noted that the latter only applies to private actors, as public authorities are subject to administrative law, which stipulates that information and documents that have been recorded in accordance with the rules of the law may not be deleted. The framework of administrative law is described in more detail above.

Purpose limitation

If, in connection with SMCS, personal data is collected for one specific purpose, it follows from Article 5(1)(b) that this data may not be used for any other purposes. Thus, data collected for the purpose of crisis management in connection with natural disasters may not subsequently be used for marketing or similar purposes.

As described above, according to Article 5(1)(a), it is also a requirement that there is a legal basis (either national or at EU level) that allows personal data to be used for the specific purpose. Alternatively, there must be valid consent from the data subjects. In the event there is a situation where consent is given, it is required that the persons are informed about the actual purpose of the collection and the possibility of them withdrawing their consent at a later stage.

Social media platforms – Terms of Services

There are countless forms of social media; blogs, wikis, photo-sharing sites, podcasts, virtual worlds, etc., used by billions of people. On a personal level, social media provides an opportunity to communicate with friends and family, learning new things, developing your interests and to be entertained.

There are also opportunities for companies and public authorities to use social media to engage in a conversation with their audience, get feedback and raise their brands.

Data tells us that approx. one third of the world uses one or more forms of social media and this number is going to increase in the coming years and that the amount of data these social media contain is something that can be used for many different purposes. The use of social media by the Belgian authorities in 2016 in

connection with a terrorist attack (Marynissen 2020) has already been mentioned, but social media can also be used for propaganda, surveillance, and the sharing of false information.

It is therefore obvious that the use of SMCS involves consideration of how the data collection is done and that the use can take place without compromising ethics and without getting into legal trouble. As there are countless social media platforms, the following will only focus on the five platforms which are broadly relevant in relation to SMCS for use in crisis management. This demarcation was therefore made based on the following criteria:

- European users (platforms such as KuaiShou are excluded)
- The number of monthly active users on the platform
- The possibility to collect data that specifically relates to crisis situations.

Please note the following represents a snapshot from November 2023 – and other dates are mentioned if they diverge from the snapshot. Both the actual functions of these platforms and their terms of services, with consequences for the legality, can change with time. Therefore it is necessary to be aware of the potential future changes.

Meta: Facebook and Instagram

Instagram and Facebook constitute two of most used as well as the fastest growing social media platforms in Denmark and will thus be particularly useful for SMCS in Denmark. Both media platforms are owned by Meta (formerly Facebook) and have the same terms of service

While Instagram is primarily used to share images (either as permanent posts or as a temporary "story" that is only visible to other users for 24 hours), the use of Facebook is much more versatile. Facebook is used to share both images and text posts. The posts can be shared both on the user's own profile and in groups where other group members can interact with the post.

On both platforms, posts can be tagged with hashtags, which are keywords used to search for other posts with the same theme.

Terms and conditions

The terms of service for the use of Meta Platforms can be found here: https://developers.facebook.com/terms.

Section 3, part a, iii of the terms and conditions on **prohibited use** states the following:

"Processing Platform Data to perform, facilitate, or provide tools for surveillance. Surveillance includes the Processing of Platform Data about people, groups, or events for law enforcement or national security purposes."

The lawful use and processing of data collected from Meta Platforms is thus dependent on the data not being used for surveillance, crime prevention or other security purposes.

In addition, the following is stated in section 3, part a, v on prohibited use:

"Placing Platform Data on, or otherwise making Platform Data available to, a search engine or directory without our prior express written consent."

In addition, the terms of service here state that *"You may not access or collect data from our products using automated methods (without our prior permission) or attempt to access data you are not authorized to access. We also reserve all our rights against text and data mining."*

Thus, it is necessary for the lawful use of data from Meta that the collection is done manually and complies with the above-mentioned restrictions. If the data is to be collected automatically, however, it is necessary that permission is obtained from Meta, that the collected data will only be available to the group that will be working with crisis management, and that the collected data is not used for alternative purposes or made available to a search engine.

Recommendation

Based on the above, it can be deduced that the most advantageous way to use data from Facebook/Instagram would be through a special permission from Meta if the data is the harvested through an automated process. If the applied method of data harvesting is manual, then there seems to be no requirement to gain a special permission from Meta and the data can be harvested, in accordance with the applicable national and international rules, see earlier sections.

The subsequent methodological consideration must focus on the fact that this collected data is not used via a proprietary database that can be used by a search engine.

X – formerly known as Twitter

X is a social media where users can follow each other and send messages to their followers. Users can choose to target messages to a specific user or make their messages public so that all users on the platform can access them. X differs from other social media in that messages are limited to a maximum of 280 characters. These are short messages, which in some cases are accompanied by images.

Messages on X are sorted with hashtags, which are keywords that are added to the message and can be searched across the platform. The medium could thus be relevant in relation to SMCS, as you could search for messages with relevant hashtags and use the information in emergency response work.

Terms and conditions

The terms of service for using X can be found here; https://help.twitter.com/en/rules-and-policies

X has been undergoing a lot of change recently with initiatives such as new features, a new name, and a number of new terms of service. This is a snapshot of the terms and conditions that apply as of September 2023.

As of September 29, 2023, the following terms and conditions apply:

(...) Misuse of the Services

(...) You may not do any of the following when accessing or using the Services (...)

(ii) probe, scan or test the vulnerability of any system or network or breach or circumvent any security or authentication measures; (iii) access or search or attempt to access or search the Services by any means (automated or otherwise) other than through our currently available published interfaces provided by us (and only in accordance with the applicable terms and conditions), unless you have been specifically allowed

to do so in a separate agreement with Twitter (NOTE: crawling the Services is permissible if done in accordance with the provisions of the robots.txt file, however, scraping the Services without the prior consent of Twitter is expressly prohibited)

Recommendation

Based on the above, it can be concluded that the lawful and appropriate use of X is based on the explicit consent of X in the context of SMCS in emergency situations, if the applied method is automated

A manual process of data harvesting is still possible within the current terms and conditions of X.

However, it should be noted that the guidelines for X can change relatively quickly, and new rules can be applied at short notice.

Reddit

Reddit differs from other social media platforms in that it encourages users to remain anonymous. Thus, users generally do not use their real name, picture, or other identifying markers on their profiles, unlike users on Meta platforms, where it is not allowed to use fake names.

Reddit is structured as an international forum with a number of sub-forums (subreddits), each focusing on a more or less specific topic. On each of these subreddits, posts can be created with text and links to either articles, images, or other online material. Posts are moderated by volunteer moderators on each subreddit and are voted up or down by the subreddit's users. Posts gain prominence on subreddits by having many upvotes and few downvotes.

In the case of particularly large events (such as war or a disaster), a "megathread" is often created and pinned to the front page of the subreddit, making it the first post users see on a subreddit, regardless of the number of upvotes.

Terms and conditions

The terms of service for the use of Reddit can be found here: https://www.redditinc.com/policies/user-agreement-april-18-2023#EEA

The terms of service on Reddit differ significantly from the terms of service of other social media. Other social media terms of service focus heavily on protecting the content and data from their platform from being accessible to third-party users. Reddit's terms of service are more focused on third parties not using content from Reddit for commercial use.

This is stated in section 3 of the Terms of Service as follows:

Except and solely to the extent such a restriction is impermissible under applicable law, you may not, without our written agreement:

- *license, sell, transfer, assign, distribute, host, or otherwise commercially exploit the Services or Content;*
- *modify, prepare derivative works of, disassemble, decompile, or reverse engineer any part of the Services or Content; or*

- *access the Services or Content in order to build a similar or competitive website, product, or service, except as permitted under any Additional Terms (as defined below).*

Recommendation

Based on the above, it can be concluded that Reddit is highly useable for SMCS in crisis situations. Data harvesting from the platform can be done by both public authorities and NGOs and can be done both manually and automatically without violating the terms of service.

Data from Reddit will also be exempt from data protection rules because of the fact that the information is not personally identifiable.

Snapchat

The social media platform Snapchat differs from other social platforms in that images and text messages are automatically deleted after viewing. However, images can also be posted in a "story" where they will be available for 24 hours. Stories can either be available only to the user's own connections on Snapchat or on the "Snap Map". The Snap Map works in such a way that all users can contribute to it and all users can access all contributions. In practice, the Snap Map is often used for larger events such as concerts, demonstrations, and the like. But it can also be used in crisis situations, such as fires, floods, or shootings.

This means you can't search for posts based on keywords or similar methods, but published stories from specific geographical areas can be easily accessed.

Terms and conditions

Terms and conditions for using Snapchat can be found here: https://snap.com/en-US/terms.

In section 7 of the terms and conditions the following is stated to be not allowed and may cause termination and suspension from Snapchat:

- *use any robot, spider, crawler, scraper, or other automated means or interface to access the Services or extract other users' information;*
- *use or develop any third-party applications that interact with the Services or other users' content or information without our written consent;*

Much like with Meta-platforms, the legality of crowdsourcing data depends on getting a written consent to do so from the owner of the service, in this case Snapchat, and that the data crowdsourcing can be done without the use of any automated means.

Recommendation

Based on the above, it can be concluded that a legal and appropriate way to use Snapchat in a crisis situation SMCS would be as following:

- **Manual** real time surveillance of areas on the Snapmap affected by a crisis.
- Get written consent from Snapchat to access and use data from Snapchat to handle crisis' and/or prevent material damage.

START HERE - When you are considering doing social media crowdsourcing (dataharvesting)

Will personal data from the posts being crowdsourced be used?

No

You do **not** have to consider data protection rules.

Is it a Danish public authority performing SMCS?

No

You do **not** have to consider administrative law

Use the guide on the next page to see how SMCS can be permitted according to platform's terms of use.

Yes

You have to consider administrative law, e.g.
- Good administrative practice

Use the guide on the next page to see how SMCS can be permitted according to platform's terms of use.

N.B: If SMCS is used to make decisions the Public Administration Act and the Freedom of Information Act must be adhered to, e.g.
- Obligation to take notes
- Obligation to journalize
- Justification
- Complaints guide

Yes

You have to consider the data protection rules, e.g.
- It must be stipulated by legislation that the use of data is allowed (the legal basis requirement)
- Only the data necessary to achieve the purpose may be used (data minimization)
- Data must not be used for purposes other than the purpose for which they were collected (purpose limitation)

Is it a Danish public authority conducting SMCS?

Yes

No

LINKS

Strengthening links between technologies and society for European disaster resilience

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 883490

Dataharvesting
Collection, categorization, sorting or anything similiar of data from platforms. The collection can be perfomed either automatically or manually.

Automatic dataharvesting
Data harvesting without human intervention, for example using algorithms or data crawlers.

Manual dataharvesting
Data harvesting with human intervention, for example, by letting employees go through a website to select and categorize relevant data.

Social media crowdsourcing (SMCS)
When dataharvesting is being done from social media it is called social media crowd sourcing.

Personal data
Personal data is any kind of information that can be related to an identifiable person - for example a name (directly identifiable), a picture (directly identifiable) or a license plate (indirectly identifiable)

Administrative law
A collective term for all the sources that together form the legal basis for public authorities. Different legal sources apply at different times depending on what the public authority is doing.
Administrative law in Denmark thus consists of the Public Administration Act, the Freedom of Information Act, decisions from authorities, court judgments, statements from the Ombudsman and a number of other sources.

Good administrative practice/Good Governance
General principles that follow from the Ombudsman's practice regarding how public authorities should act.

Decisions
Decisions differ from actual administrative activities and procedural decisions, as decisions aim to clarify the legal reality. Decisions are unilateral communications from a public authority to either one or more citizens, which aim to establish what the applicable law is in a specific situation

Instagram and Facebook (Meta platforms)
SMCS is permitted if it is performed manually, as long as the data is not used for monitoring purposes and as long as the data is not made available to a search engine.

If SMCS is to be performed automatically, a separate agreement must be made with Meta Inc..

X - formerly known as Twitter
SMCS is only permitted if an agreement is made with X Holding Corp. regardless of whether SMCS is manual or automatic

Snapchat
SMCS is permitted if performed manually.

If SMCS is to be performed automatically, a separate agreement must be made with Snapchat Inc.

Reddit
SMCS is permitted, whether automatically or manually, as long as the data is not used for commercial purposes.