Kent Academic Repository

Full text document (pdf)

Citation for published version

Aydn, Kübra, Salam, Rahime Belen, Li, Shujun and Bülbül, Abdullah (2020) When GDPR Meets CRAs (Credit Reference Agencies): Looking through the Lens of Twitter. In: Proceedings of the 13th International Conference on Security of Information and Networks. . 16:1-16:8. ACM ISBN 978-1-4503-8751-4.

DOI

https://doi.org/10.1145/3433174.3433586

Link to record in KAR

https://kar.kent.ac.uk/89476/

Document Version

Author's Accepted Manuscript

Copyright & reuse
Content in the Kent Academic Repository is made available for research purposes. Unless otherwise stated all content is protected by copyright and in the absence of an open licence (eg Creative Commons), permissions for further reuse of content should be sought from the publisher, author or other copyright holder.

Versions of research
The version in the Kent Academic Repository may differ from the final published version.
Users are advised to check **http://kar.kent.ac.uk** for the status of the paper. **Users should always cite the published version of record.**

Enquiries
For any further enquiries regarding the licence status of this document, please contact:
researchsupport@kent.ac.uk

If you believe this document infringes copyright then please contact the KAR admin team with the take-down information provided at **http://kar.kent.ac.uk/contact.html**

KAR Kent Academic Repository

University of Kent

When GDPR Meets CRAs (Credit Reference Agencies): Looking through the Lens of Twitter

Kübra Aydın*
175101125@ybu.edu.tr
Ankara Yıldırım Beyazit University
Ankara, Turkey

Rahime Belen Sağlam
R.Belen-Saglam-724@kent.ac.uk
University of Kent
Canterbury, UK

Shujun Li
S.J.Li@kent.ac.uk
University of Kent
Canterbury, UK

Abdullah Bülbül
abulbul@ybu.edu.tr
Ankara Yıldırım Beyazit University
Ankara, Turkey

ABSTRACT

Collecting information about consumers and businesses from various sources, Credit reference agencies (CRAs) help many organizations such as financial institutions to assess creditworthiness of applicants and customers of their services. CRAs' business model depends on processing a high volume of personal data including highly sensitive ones, which must be processed within the relevant legal frameworks in different countries they operate their business, e.g., the European Union's new GDPR (General Data Protection Regulation). This paper reports a data-driven analysis of CRA- and GDPR-related discussions on Twitter. Our analysis covers the three largest multi-national CRAs: Equifax, Experian and TransUnion and we also looked at the UK's data protection authority, ICO, and two UK-based privacy-advocating NGOs, Privacy International and Open Rights Group (ORG). We have analyzed public tweets of their official Twitter accounts and other public tweets talking about them. Our analysis revealed a very surprising lack of awareness of CRA- and GDPR-related data privacy issues within the general public and an astonishing lack of active communications of CRAs to the general public on relevant GDPR-related privacy issues: out of 39,549 collected tweets we identified only 153 relevant tweets (0.387%). This small number of tweets are dominated by mentions of security issues (%73.2), especially data breaches affecting CRAs, not data subject rights or privacy issues directly. Other tweets are mainly about complaints regarding inaccurate data in credit files and questions about how to exercise right to rectification, just two of many data subject rights defined in the GDPR.

KEYWORDS

credit reference agencies, general data protection regulation, GDPR, online social networks, social media, twitter, data protection, law, privacy, transparency, communication

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SINCONF'20, November 4–7, 2020, Istanbul, Turkey
© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-8751-4…$15.00
https://doi.org/10.1145/3433174.3433586

ACM Reference Format:
Kübra Aydın, Rahime Belen Sağlam, Shujun Li, and Abdullah Bülbül. 2020. When GDPR Meets CRAs (Credit Reference Agencies): Looking through the Lens of Twitter. In *Proceedings of 13th International Conference on Security of Information and Networks (SINCONF'20).* ACM, New York, NY, USA, Article 16, 8 pages. https://doi.org/10.1145/3433174.3433586

1 INTRODUCTION

Credit scoring is a technology used as a decision-making tool to measure the creditworthiness of applicants (e.g., opening a bank account, applying for a loan from a lender, or trying to rent a property). In order to measure the credit score of an individual, a range of information is taken into account including information provided directly by the individual whose credit is being evaluated. It is often the case that all or part of such information is provided to the service provider by an intermediate agency called a credit reference agency (CRA). CRAs are informational brokers that collect information from various sources and provide credit-related information on individual and corporate consumers to other organizations to facilitate contract establishments and transactions between them. Different terms are used to refer to CRAs in different countries and sectors, e.g., "credit bureau", "consumer credit (reporting) agency" and "consumer credit information supplier". We use the term CRA throughout the paper because it is used by the Association of Consumer Credit Information Suppliers (ACCIS)[1], an industrial association of CRAs in Europe, and in the UK Consumer Credit Act 1974[2].

The sensitive nature of personal data CRAs collect from many sources and store for serving their customers has profound privacy implications. Although CRAs do proactively disclose what data they collect and how they process the data, e.g., the three main CRAs in the UK (Equifax[3], Experian[4] and TransUnion[5]) via their joint Credit Reference Agency Information Notice (CRAIN) [7], many data protection issues are not widely known by the general public. For instance, CRAs do not depend on data subjects' explicit consents to collect and process personal data, but other "lawful

[1] https://accis.eu/
[2] https://www.legislation.gov.uk/ukpga/1974/39/part/X
[3] https://www.equifax.co.uk/
[4] https://www.experian.co.uk/
[5] https://www.transunion.co.uk/

bases" allowed by data protection laws including the latest European Union (EU)'s General Data Protection Regulation (GDPR)[6], especially the so-called "legitimate interests" (see Section 3 of [7] and also the guideline on CRAs from the ICO, UK's data protection authority [12]). In addition to data protection laws, CRAs are normally registered with financial regulators and their activities regulated by consumer credit legislation (e.g., Consumer Credit Act 1974 in the UK [22]).

Despite the legal grounds of CRAs' data collection and processing practices, privacy concerns on CRAs and the strengthened data subject rights introduced in the GDPR in 2018 led some privacy-advocating non-governmental organizations (NGOs) to challenge the legal grounds of CRAs. In one of those campaigns, Privacy International[7], a London-based NGO, called people to request their data deleted by seven those data brokers including two large CRAs (Equifax and Experian) and files official complaints to the ICO requesting an assessment notice on those data brokers' compliance with the EU GDPR and the UK Data Protection Act 2018 [16–18]. They have started another campaign 5 months after GDPR came into effect and highlighted rights of individuals with regard to the protection of their data strengthened by the GDPR [1].

Despite the sensitive nature of data collected and processed by CRAs and the campaigns organized by privacy-advocating NGOs, we noticed a lack of research on how the general public perceive the privacy debate around CRAs. To fill this research gap, we conducted a data-driven analysis of CRA- and GDPR-related discussions on Twitter. We focused on the GDPR as a representative data protection law because it covers many countries and has triggered a lot of discussions on data subject rights in other application areas. We expected that there should be a lot of relevant discussions on Twitter about both CRAs and the GDPR. As a matter of fact, in its complaint letter sent to the ICO in 2018 [18] after the GDPR became effective, Privacy International heavily cited data subject rights defined in the GDPR.

Our analysis is based on tweets of official accounts of six organizations, the three largest multi-national CRAs operating in the UK and many other countries, Experian, Equifax and TransUnion, the UK data protection authority ICO[8], and two privacy-advocating NGOs, Privacy International and Open Rights Group (ORG)[9]. We collected tweets from official accounts of the six organizations and also public tweets mentioning the name or their official Twitter accounts, and analyzed their contents.

Our work revealed a very surprising lack of awareness of CRA-related data privacy issues within the general public and an astonishing lack of active communications of CRAs to the general public on relevant privacy issues: out of 39,549 collected tweets we identified only 153 tweets (0.387%) mentioning both CRAs and the GDPR. Most tweets in this small set are discussions on security issues (%73.2), especially data breaches affecting CRAs, not on data subject rights or privacy issues. The remaining tweets are mostly about complaints regarding inaccurate data in credit files and questions about how to exercise right to rectification, just two of many data subject rights defined in the GDPR.

[6]https://eur-lex.europa.eu/eli/reg/2016/679/oj
[7]https://privacyinternational.org/
[8]https://ico.org.uk/
[9]https://www.openrightsgroup.org/

The rest of the paper is organized as follows. In the next section, we give an overview of both CRAs and the GDPR as useful background. Section 3 introduces related work concerning both CRAs and GDPR. After that we introduce data we used for our study in Section 4 and the results of our data analysis in Section 5. The final section concludes the paper.

2 BACKGROUND

2.1 CRAs

CRAs as a sector have tried to make their data collection and processing practices transparent to the general public, and they have published a number of reports [2, 7, 19]. Other organizations especially the national data protection authorities (DPAs) and consumer organizations also have interests in how CRAs operate and have published their own guidance documents or reports [11, 21]. In this section, we give a brief overview of CRAs' data collection and processing practices based on the above public reports and guidance documents.

In terms of organizations providing data to CRAs, they include many sector such as banks (public, private or postal and cooperative banks and credit unions), leasing companies, credit card suppliers, mortgage providers, retail credit suppliers, insurance companies, debt collectors, 'enforcement divisions' (courts, tax authorities and the police), governmental departments, telecommunication operators, internet service providers, television suppliers, utility (electricity, gas and other fuel) suppliers and brokers. Information collected by CRAs can be categorizes into a number of classes, e.g., general consumer data (e.g., personally identifiable information), information on the credit applications, legal information, information on family, income, assets, details on credit contracts, loan data, and payment data.

Given the value of the information collected by CRAs and the need to make decisions based on accurate and up-to-date information, source organizations update CRAs regularly about any changes of such information, and often electronically in real time. CRAs normally employ independent inspection and internal controls such as periodical checks as data quality strategies. In addition to those internal checks, CRAs also respond directly to consumer complaints and data protection/regulator requests from organizations and the authorities. Individuals also have the right to make a complaint to the national data protection authority, if there is an obvious inaccuracy and neither the CRA and nor actual data source is willing to correct [13]. CRAs have to comply with various laws and regulations both at the regional (e.g., EU) and the national level. For instance, the EU GDPR and the UK Consumer Credit Act 1974 give individuals the right to request changes to inaccurate data held by the CRAs, although this may have to be done by contacting the relevant information source organizations. They have to align with national laws on the protection of personal data and consumer protection and banking laws to specific credit reporting [19]. Principal national regulator differs from country to country where it could be data protection authorities, the Ministry of Finance, Parliament or National Bank [19].

With the rich information collected, CRAs can serve other organizations in both the private and public sectors, who can request access to relevant information relating to an applicant of their service

and decide if the application should be accepted. Such information consuming organizations normally establish their legal ground to access such information by obtaining the data subject's consent. It is important to note that an information receiving organization can access the information from CRAs only when it has permissible reason as defined in the law. It is made available on request to customers of the CRA for the purposes of credit risk assessment, credit scoring, or for other purposes such as employment consideration or leasing an apartment within a regulatory framework [9]. There are several possible ways of gathering data subject's consent: explicit informed consent, unambiguous informed consent as part of general terms and conditions, and through a specific agreement signed by the data subject. However, some reports revealed that some negative data about individuals could be collected without any form of consent.

Some reports also cover privacy concerns raised by consumers. Such concerns were mostly about the collection and sharing of credit data in general. Other privacy concerns include difficulties for data subjects to access their own data for rectification and the lack of transparency concerning third-party access to their credit data.

2.2 GDPR

The GDPR, which became enforceable across the whole EU in May 2018, is the European Commission's most recent attempt to protect the privacy of data subjects in the EU (not just EU citizens) regardless of the location of their data, and personal data of any data subjects (not necessarily present in the EU) that are collected or processed in the EU. In order to achieve its set goals, the GDPR defines a set of principles and the lawful basis for data controllers and data processors to process personal information. It also provides several rights for data subjects, individuals whose data are collected and processed. Data controllers are defined in Article 4 as natural or legal persons, public authorities, agencies or other bodies who dictate how and why data is going to be used by the organization and data processors processes any data that a data controller gives them. Given these definitions, a CRA could be considered both as data processor and a data controller.

In this section, we present some of the most pertinent elements of the GDPR, to facilitate the later discussions on our work and results. In Article 5, the GDPR defines several data protection principles an entity must comply with when processing personal data. The first set of principles requires data controllers to process personal data lawfully, fairly and in a transparent way. The data minimization requires controllers to ensure that personal data processed is adequate, relevant and limited in relation to the purposes for which they are processed. Another principle,storage limitation, dictates personal data to be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed. The accountability principle requires data controllers/processors to take responsibility for what they do with personal data and how they comply with the other principles. Those bodies are also expected to have appropriate measures and records in place to be able to demonstrate their compliance. Another responsibility of data controller is given in Article 34(1) which requires data controllers to notify identified individuals impacted by the breach in the event of a security breach. Organizations are supposed to report data breaches within 72 hours.

In addition to those principles, GDPR also outlines lawful basis for processing personal information which includes obtaining explicit consent and legitimate interest. If a data controller/processor has legitimate interest and can show that the processing is necessary to achieve it, they do not need to obtain consent. Finally, the GDPR provides several rights to individuals. The right to be informed lets individuals know what is being done with their information, and right to access allows data subjects to ask for a copy of their personal data, the purposes of processing their data, the categories of the data being processed, and the third parties or categories of third parties that will receive their data. Under the GDPR, data subjects have also a right to rectify or erase inaccurate or out of date information which is known as right to rectification. The GDPR also introduces a right for individuals to have personal data erased if there is no longer a lawful basis for processing or if the data subject withdraws consent.

3 RELATED WORK

3.1 CRAs' Responses to GDPR

ACCIS published an article in February 2020 [3] to summarize their contributions on informing the European Commission for the latter's 2020 evaluation report on the application of the GDPR. In its article, ACCIS claimed that its members (CRAs in Europe) had taken GDPR readiness very seriously. Updating data protection statements in order to comply with the extended transparency obligations under the GDPR is given as one of the actions taken by CRAs. The ACCIS article states that such statements have been made visible via CRAs' websites, and links to those statements have been covered in all email communications. In addition, online portals were set up for this purpose. Launching industry-wide public information notices is given as another strategy to prove transparency about the type of information CRAs hold and the legality of handling personal data. The ACCIS article also talks about statistics of GDPR related data subject requests and how CRAs work with national data protection authorities to ensure the GDPR compliance. It also touches how the GDPR could cause challenges on applications of new technologies and on data breach notifications.

Communicating with several CRAs, ACCIS reported that there was an overall increase in the number of data subject requests including full subject access requests and rectification requests after the GDPR. Data erasure requests and data portability requests were much rare. ASSICS also stated that complaints and court action usually relate to the right of access, the right to erasure and the right to object. It underlines the importance of ongoing dialogue with DPAs for possible situations of non-compliance. On the other hand, inconsistencies among DPAs is also highlighted and it is added that when there is a legal uncertainty, CRAs would turn to their local DPAs for guidance. The report explicitly states that GDPR has made it not easier to enable new technologies and methods. Regarding the developments of new technologies such as artificial intelligence, CRAs are reported to have different opinions where some CRAs think that current safeguards are sufficient to grant fair and trustworthy data processing whereas others think that additional rules

or guidance are needed to clarify how the data protection principles should be applied in practice. Regarding the data breaches, ACCIS highlighted the risk that controllers might overreact when it comes to notifying a data breach due to unbearable fines. It argues that this may lead companies notify in circumstances where notification was not essential thereby putting unnecessary strain on DPAs.

While the ACCIS article [3] highlighted CRAs' efforts to be GDPR ready and compliant, the Credit Reference Information Notice (CRAIN) Version 1.1 [7], jointly compiled by three large CRAs in the UK (Equifax, Experian and TransUnion), has no a single mention of the GDPR, even though it was revised in March 2020, nearly two years after the GDPR became effective. This phenomenon also appeared in our results of analysis on tweets from the three CRAs' official accounts (see Section 5 for more details), reflecting a lack of active communication by CRAs to the general public on GDPR issues.

3.2 GDPR-related Discussions on Social Media

Considering the massive amount of studies that analyze the impact of the GDPR in several contexts such as cloud computing, Internet of Things and blockchain technologies, there has been surprisingly less research looking at social media discussions regarding the challenges raised by the GDPR. In one of those studies, Yang et al. investigated the discussions on Twitter around the right to be forgotten (RtbF) through social network analysis [24]. Researchers examined the various topics discussed in relation to RtbF and the role of influencers in this debate. Google's role in RtbF and Russia (discussions in Russia regarding how public figures could abuse the law by removing compromising reports to enhance their online reputation) have been reported as the dominating topics around this right. On the other hand, companies that are affected by the RtbF decision directly, experts including privacy researchers and lawyers, and news portals have been identified to be the key players in the network for this discussion [24].

In a similar study, Grudz et al. studied the discussion about the GDPR in a broad sense with the aim of examining public opinions and organizational public relations strategies about the GDPR [10]. Researchers collected tweets with the hashtag #gdpr and applied social network analysis to find opinion leaders. Their results indicate that the GDPR is being actively discussed by a variety of stakeholders, but especially by cyber-security and IT-related firms and consultants. One of the most interesting finding has been reported as that some of the stakeholders that were expected to have a more active role were less involved, including companies that store or process personal data, government and regulatory bodies, mainstream media, and academics [10].

Social media analysis on data privacy is not limited to discussions around the GDPR. There are also studies conducted to understand what people think and how they are affected as customers from data breaches. After the discovery that Facebook gave unauthorized access to personally identifiable information (PII) of more than 50 million Facebook users to the data firm Cambridge Analytica [5], Gonzalez et al. conducted a study to understand how data privacy concerns vary across the world. They conducted a cross-language study of online conversations to compare how people speaking different languages (English vs Spanish) react to data privacy breaches. Revealing the cultural differences in data privacy perspectives, their results show that the tweets written in English, in which Americans were the most active group of users, blame companies about these data breach, whereas Spanish speakers are more likely to blame users. In a similar study, Vemprala and Dietrich tried to cover all the discussions regarding data breaches and collected tweets that contain predefined keywords related to data breaches [23]. They mainly focused on information diffusion and tried to identify what characteristics of a breach message makes the information diffuse, whether the presence of evidence in the form of media files, URLs and videos, drive information diffusion and finally who among the social media users exchange wide amounts of breach messages. Analyzing the messages in both linguistic and visual perspective through social networks, researchers found that the messages that involve the technicalities, threat and severity related security characteristics spread fast. The messages are spread widely by technology groups and around groups which are doing studies about security instead of conventional news channels [23].

While there have been quite some research on social media discussions regarding data privacy issues, there has been very little research on discussions about CRAs. In one of such studies, Novak and Vilceanu [15] analyzed crisis communications via Twitter after the 2017 Equifax data breach, which saw 143 million US customers of Equifax had information compromised [4]. They collected tweets with some predefined hashtags (#Equifax, #Equifail, #WheresMyData, #Equifaxbreach, etc.) during the 3 weeks after the breach was disclosed by Equifax. Through a qualitative analysis, they identified discourses that represented user frustrations and reactions to the incident. Breaking news, anger and outrage, and blame attribution were reported as major discursive patterns.

The aim of our work is to explore all privacy concerns observing the interaction between CRAs and their customers. Concerns raised by different groups such as national authorities in charge of data protection issues are also covered to enrich our findings.

4 METHODOLOGY AND DATA USED

In this section we explain the methodology and data we used in our work, including how we collected raw data and how we conducted pre-processing and manual selection of tweets for further analysis.

4.1 Organizations and Social Media Accounts Studied

The three largest CRAs (Experian, Equifax, and TransUnion) [14] were selected as representative CRAs for our study. Collectively they hold credit records of billions of individuals and businesses. In addition to the three CRAs, we also looked at the UK's data protection authority, ICO, and two UK-based privacy-advocating NGOs, Privacy International and Open Rights Group (ORG).

For social media platforms, we decided to focus on Twitter because it is more open and many past studies on social media used data on Twitter. Some organizations have multiple Twitter accounts, and we decided to limit our study to the main official account, assuming that one has a broader coverage on issues such as the GDPR and faces all visitors of the organization. The actual Twitter accounts used can be found in Table 1.

4.2 Raw Twitter Data

Tweets used in the study were collected via a Python library called Tweepy[10] that allowed us to access different Twitter APIs more easily. Firstly, we used Twitter's Search API to collect public tweets that cover the names of CRAs. We started collecting the data in November 2019 and kept crawling for 3 months to obtain representative samples. Secondly, we collected tweets in which official twitter accounts of the six target organizations were mentioned using ScrapeHero Cloud[11]. The tool allowed us to retrieve tweets published since January 2018 until February 2020 in which the data collection process was completed. Lastly, we collected the most recent tweets published by official Twitter accounts of the six target organizations using the User_Timeline API which provided tweets of Experian (since October 2019) and of Equifax (since May 2019). The number of tweets posted by TransUnion was relatively low, so we retrieved tweets of this CRA since March 2017. It was a similar case for the ORG where we could access to tweets posted since April 2017. We could access to tweets of the ICO since February 2019 and tweets of the Privacy International since May 2018. At the end, we collected 39,549 tweets in total. The number of tweets crawled for each group can be seen in Table 1.

Table 1: The distribution of tweets in our dataset: A = tweets collected from each target organization's official Twitter account(s), B = public tweets mentioning each target organization's official Twitter account(s), C = public tweets mentioning each target organization's name or acronym

Organization (Account)	#(A)	#(B)	#(C)
Experian (@Experian)	3,400	540	13,629
Equifax (@Equifax)	3,598	210	2,574
TransUnion (@TransUnion)	3,330	120	1,878
ICO (@ICOnews)	3,292	233	-
Privacy International (@privacyint)	3,268	128	-
ORG (@OpenRightsGroup)	3,244	105	-

4.3 Data Pre-Processing and Cleaning

Our data collection protocol includes the following three main steps.

1) **CRA-related filtering**: We crawled the tweets posted by the 3 CRAs, and tweets of the other 3 organizations that mentioned those 3 CRAs. This allowed us to narrow down our focus to the CRA context.

2) **GDPR-related filtering**: As we obtained the tweets of the 6 target organizations, in order to eliminate irrelevant tweets, we identified a list of keywords considering related elements of the GDPR in the CRA context. We extracted 85 keywords manually from the GDPR document itself and the ICO's guideline[12] to the GDPR (see Table 2). We limited our analysis to tweets that cover at least one of the identified keywords. We ended up with more than 12,000 tweets after this filtering step.

3) **Manual Filtering**: A majority of the tweets obtained in the second step are generic tweets published to inform people about data privacy or the GDPR. Since the aim of the study is to observe meaningful discussions around both CRA- and GDPR-related discussions, we included tweets that have direct and implicit mentions of the GDPR or any element of it. We also eliminated informative tweets manually that give very general information about the GDPR. After applying this step, we surprisingly found that only a very small set of 153 tweets remained as relevant, which were used for our further analysis.

Table 2: Keywords

Access	Accountability	Accuracy
Adequate	Automated Decision	Certification
Codes of conduct	Confidential	Consent
Contract	Controller	Correct
Criminal	Cyber	Cyberattack
Dataintegrity	Dataleak	Datasafety
Delete	Encryption	Erase
Erasure	Fair	Fairness
Forget	Forgotten	Format
GDPR	Hold	Impact
Inform	Law	Lawful
Legal	Legitimate	Limit
Limitation	Long	Machine Readable
Minimisation	Minimum	Needed
Object	Obligation	Offence
Outside	PECR	Period
Personal	Probability	Principle
Processing	Processor	Profiling
Protection	Public Task	Purpose
Rectification	Relevant	Remove
Request	Restrict	Retain
Retention	Personal	Revoke
Right	Secure	Security
Sell	Sensitive	Sold
Special Category	Storage	Third Party
Third Parties	Transfer	Transmit
Transparency	Transparent	Update
Vital Interest	Confidentiality	Withdraw
Years		

4.4 Analysis

We manually labeled the 153 tweets to uncover CRA-related data privacy issues discussed on Twitter. While labelling, we took the definitions given in the GDPR regarding individual rights, lawful

[10] https://www.tweepy.org/
[11] https://www.scrapehero.com/
[12] https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/

basis for processing and key principles into consideration and performed the labelling accordingly. The labels were checked by two independent researchers. The identified labels and their distributions are given in Figure 1.

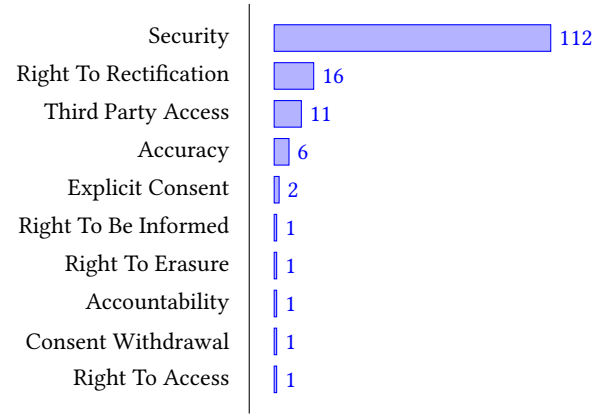


Figure 1: Distributions of tweets per GDPR-related topic

5 RESULTS

Reviewing the 153 relevant tweets, it is clear that data privacy discussion on Twitter were dominated by data breach incidents. Sharing personal data with third parties, other generic security issues and accuracy of the data processed by CRAs are the other popular topics. Although not very common, we can still see tweets regarding data subject rights including the right to be informed, the right to erasure and the right to rectification.

We also observed that discussions were dominated by individuals (109 tweets out of 153), while CRAs published data privacy related tweets only to inform their followers regarding their security measures or to reply their customers' enquiries about inaccurate data in credit files. Surprisingly, we could not find any tweets published by the ICO, Privacy International or Open Rights Group, which explicitly mentioned both CRA- and GDPR-related topics.

In order to identify cultural differences in reactions to CRAs, we have also tried to access country information of the individuals. Among the 86 unique accounts who involved in the discussions, we could have accessed country information of 56 accounts. A majority of the individuals (44) are from the US (51.2%) and there are only 5 accounts from the UK (5.8%). The other countries were the Netherlands, South Africa, and Algeria.

5.1 Security

It is obvious that security issues dominate the discussions on GDPR-related discussions in CRA context. There are two main security issues; one discussed by individuals and the one by CRAs. Majority of the tweets published by individuals are about data breach incidents whereas CRAs' security discussions are limited to information they provide regarding their security measures.

5.1.1 Data Breach Incidents. Among the 153 tweets, 72 of them have been identified to be related to data breaches incidents. Key role players in those discussions are individuals where Equifax published 3 and Experian published 2 tweets. The tweets published by Equifax are informative tweets given as a reply to individuals about 2017 data breach. An example can be given as;

> "@DrFairkid Thank you for contacting Equifax. If you have a question about the 2017 data breach settlement, please visit: https://t.co/EVlbzHwHQo. -George R."

Two tweets published by Experian provide links to podcast published on the company's website about Experian data breach resolution.

We have observed several tweets that reminds the deadline for claims for cash payment and free services for the ones who are a part of this breach. There also complaints about this application such as:

> "What's the security of my private data worth? Apparently, less than $9. Thanks, @Experian!"

It is possible to add that tweets under this category are the ones that cover anger or blame attribution the most. Following tweets can be given as examples;

> "@Equifax So you loose my data, take 6 months to tell me, fail to respond within time limit, admit loosing data but dont think I have a complaint to uphold? WTF are you idiots? Complaint lodged with UK Information Commissioner, see you in Court!"

5.1.2 General Security Issues. This is the topic where CRAs dominated the discussions by introducing their security measures. A tweet published by TransUnion in 2018 November can be given as a representative example to this category:

> "We're proud of the credit and identity protection tools that we provide to consumers. Read our Corporate Responsibility Report at: ..."

or another one published by Equifax in August 2018

> "We've been working diligently to improve cybersecurity efforts. Learn more about how our Workforce Solutions team is transforming security: ..".

CRAs also use Twitter to suggest their customers to freeze their credit to avoid potential identity theft. Examples can be seen as follows;

> "The idea behind a freeze is that you are blocking access to your report from a potential identity thief...A credit freeze may seem like an extreme move, but it can be a powerful way to protect against identity theft."

which was published by Experian in December 2019 and

> "You protect your smartphone, so you should protect your credit too. Freeze your credit and keep identity thieves out - for free!."

which was published by TransUnion in August 2019 respectively. Here it may worth to note that a credit freeze means taking control on financial information given in your credit report. This prevents identity theft and block access of potential creditors to credit reports [20]. After freezing a credit, only current creditor and government agencies can access the credit report in some emergence cases [8]. Setting up an fraud alert to warn the creditor in case of identity theft or any other fraud is another approach suggested by CRAs [6].

Very few tweets published by individuals are complaints regarding data confidentiality where an interesting example can be given as:

> @Equifax @AskEquifax my account currently showing other people's financial information. Insane level of privacy breach. Your reset password page also does not work. Any UK phone contacts?".
> Very concerning that @ClearScore and @Equifax @AskEquifax have knowingly allowed me to have access to my neighbours personal and financial data via my credit report. Only response I've had from them is "it happens" #GDPR #gdprwho?

It is possible to report privacy concerns of the individuals as a consequence of existing or potential data breaches.

> Chinese Hacking Is Alarming. So Are Data Brokers. Companies like Equifax threaten our personal privacy and our national security".

5.2 Right to Rectification

We have identified 16 tweets regarding right to rectification and 11 of them were published by CRAs to engage with their customers due to an inaccurate information in credit files. It is surprising to see that CRAs reply to complaints via Twitter and ask their customer details via direct messages to response their enquiries. Following tweet of TransUnion can be given as a representative example:

> "@ShalomStephens Hi Shalom! I'm sorry to hear you are now seeing this inaccuracy. We can assist with correcting this information on your TransUnion credit file...!"

which was sent as a reply to a tweet of a customer:

> "Please explain to me @Experian & @TransUnion how can I have a 30 day late payment reported to my credit report for an account that was paid in full over a month ago? ..."

A reply from Equifax to a similar user tweet can be given as:

> "@Ebony07169884 We apologize for the trouble this has caused, please send us a DM with your personal information, so we can properly assist you."

5.3 Accuracy

Under this category, we have limited our focus to tweets that only disclose/mention data inaccuracies without asking for correction. Following this approach we have identified 7 tweets 6 of which were published by individuals. The most interesting tweet can be given as follows:

> "@Equifax has been hit with a class action lawsuit from a consumer who claims she is a victim of inaccurate credit reporting..."
>
> "GDPR on the credit agencies? They hold inaccurate info on you. They must correct it"
>
> "@Equifax .. I raised this with you weeks ago and it is still not resolved. Your colleagues are not even replying to me or updating me anymore so I may just have to report to the ICO as your company clearly doesn't understand its GDPR responsibilities fully."

5.4 Third Parties

GDPR defines "third party" in Article 4 as a natural or legal person, public authority, agency or body other than the data subject, controller, processor and persons who, under the direct authority of the controller or processor, are authorized to process personal data. We have identified 11 tweets under this category all of which were published by individuals to display their frustrations and reactions to third party data access. Some examples can be seen below;

> "@Experian Stop selling customer information to companies. Since making an account with you all 5 days, I:.ve been receiving at least 8 spam emails a day."
>
> "@TransUnion This was just after your employees stole their IDs and sold them online and to your business partners aka collection companies!."
>
> "So turns out @Experian sold my contact info to credit card companies without my consent ... Pretty underhanded. Time for some subject access requests I think. #GDPR"

5.5 Right to be Informed

Even though right to be informed has stated to be one of the rights that challenges CRAs the most [3], we have identified cery few tweets related to this right;

> "#fact 88% of consumers want more control over the use of their #data. Find more about #consumerpreferences for #security #personalisedexperience and #digitalengagement in the 2020 Global Identity & Fraud Report. https://t.co/CyMPtIRsu8".
>
> "..Equifax are sending a data processor to your system that gathers personal data and returns an automated decision. Do you inform users of this? How does GDPR apply for cases like this?".

5.6 Right to Erasure (Right to be Forgotten)

We could have identified only one tweet regarding right to be forgotten which points to the impossibility of exercising this right.

> "@wbm312 At least Transunion is candid about selling your information. Even it you're unable to get the info deleted, a little more power over non-transaction use of your data is huge."

5.7 Accountability

Our dataset covers just one tweet about accountability which has been published by an individual to share his/her reaction regarding the responsibility of the CRAs.

> "@Equifax you need to be accountable to each persons data stolen every person that this affects remove all inquiries and increase scores by 50points PERIOD this pisses me off https://t.co/KRzgQzVfht"

5.8 Consent

Given the fact that the CRAs do not need to collect consent from their users, we were expecting to see some data privacy discussions around this issue. However, we have observed only 3 tweets mentioning consent.

> "You can opt out of the sale of your data (as well as request access and deletion) at @TransUnion..."

One of those tweets was about the right to withdraw consent regarding emailing service of Experian.

" @Experian stop emailing me. I've unsubscribed #GDPR"

6 CONCLUSION AND FUTURE WORK

We have conducted a study based on a large database of the public communications on Twitter made by CRAs, their followers (or others who mention them in their tweets) and organizations that organise campaigns against CRAs. Even though it is widely known that CRAs collect sensitive personal information from several sources and process it, to the best of our knowledge, this is the first study that investigates the privacy discussions regarding this processing. Our study reveals that CRAs prefer to be salient about data privacy discussions from GDPR's point of view. Perhaps most interestingly, organisations that organise data privacy campaigns or even ICO were not involved in the privacy discussions on Twitter either. Among the 153 related tweets, 109 of them were published by individuals which can be interpreted as data privacy discussion on Twitter in CRA context are dominated by complaints or questions of individuals. Data breaches, inaccurate personal information in credit files and the difficulty to correct them, finally third party data access have been observed as the main privacy concerns of individuals.Considering the limitation of discussions on Twitter, we call for urgent more research into the interfaces between data protection law and the CRAs.

Since the ultimate goal of this study is to understand CRA- and GDPR-related discussions on Twitter, our data collection step might have led us to miss privacy discussions that lack explicit mentions of the GDPR or CRA elements. Our focus on only 3 large CRAs may have led to a lack of coverage of discussions on other CRAs. In our future work, we plan to broaden our scope and investigate more CRAs and look into tweets covering implicit discussions on the GDPR and CRAs and also general discussions on CRA-related privacy issues beyond the GDPR. We also plan to study CRA accounts from different countries to better understand cultural differences behind related discussions.

REFERENCES

[1] Our complaints against acxiom, criteo, equifax, experian, oracle, quantcast, tapad | privacy international. https://privacyinternational.org/advocacy/2426/our-complaints-against-acxiom-criteo-equifax-experian-oracle-quantcast-tapad

[2] Association of Consumer Credit Information Suppliers (ACCIS): ACCIS 2017 survey of members: Analysis of credit reporting in Europe. https://accis.eu/wp-content/uploads/2018/11/AUG18_ACCIS-Survey-of-Members-2017_FINAL.pdf (2018)

[3] Association of Consumer Credit Information Suppliers (ACCIS): ACCIS' contribution to inform the preparation of the European Commission's evaluation report of May 2020 on the application of the GDPR. https://accis.eu/wp-content/uploads/2020/02/FINAL-External-facing-ACCIS-Contribution-2020-Evaluation-GDPR-Feb-2020.pdf (2020)

[4] BBC News: Massive Equifax data breach hits 143 million. https://www.bbc.co.uk/news/business-41192163 (2017)

[5] Cadwalladr, C., Graham-Harrison, E.: Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. News report, The Guardian, https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election (2018)

[6] Equifax, Inc.: 7 things to know about fraud alerts. https://www.equifax.com/personal/education/identity-theft/7-things-to-know-about-fraud-alerts/

[7] Equifax Limited, Experian Limited, TransUnion International UK Limited: Credit Reference Agency Information Notice (CRAIN). Version 1.1, https://www.equifax.co.uk/crain/ (2020)

[8] Federal Trade Commission, US: Credit freeze FAQs | FTC consumer information. https://www.consumer.ftc.gov/articles/0497-credit-freeze-faqs

[9] Gichimu, S.: Credit Reference Bureaus, Loans Advancement and Recovery Performance by the Higher Education Loans Board of Kenya. Ph.D. thesis, University of Nairobi, Kenya (2013)

[10] Gruzd, A., Abul-Fottouh, D., Mashatan, A.: Who is influencing the #GDPR discussion on Twitter: Implications for public relations. In: Proceedings of 53rd Hawaii International Conference on System Sciences (2020)

[11] ICO, Information Comissioner's Office, UK: Credit. https://ico.org.uk/your-data-matters/credit/

[12] ICO, Information Comissioner's Office, UK: Do the CRAs need my consent to hold all this information on me? https://ico.org.uk/your-data-matters/credit/#consent

[13] ICO, Information Comissioner's Office, UK: What should i do if my credit file is inaccurate? https://ico.org.uk/your-data-matters/credit/#file

[14] Irby, L.: The 3 major credit reporting agencies and what they do: Credit bureaus collect information about your creditworthiness. https://www.thebalance.com/who-are-the-three-major-credit-bureaus-960416 (2020)

[15] Novak, A.N., Vilceanu, M.O.: "the internet is not pleased": Twitter and the 2017 Equifax data breach. The Communication Review **22**(3), 196–221 (2019)

[16] Privacy International: Tell companies to stop exploiting your data! https://privacyinternational.org/campaigns/take-control-your-data

[17] Privacy International: Our complaints against Acxiom, Criteo, Equifax, Experian, Oracle, Quantcast, Tapad. https://privacyinternational.org/advocacy/2426/our-complaints-against-acxiom-criteo-equifax-experian-oracle-quantcast-tapad (2018)

[18] Privacy International: Request for an assessment notice of data brokers: Experian & Equifax (the credit reference 'data brokers'). https://privacyinternational.org/sites/default/files/2018-11/08.11.18%20Final%20Complaint%20Experian%20%26%20Equifax.pdf (2018)

[19] Rothemund, M., Gerhardt, M.: The European credit information landscape: An analysis of a survey of credit bureaus in Europe. European Credit Research Institute (ECRI) Industry Survey, commissioned by Association of Consumer Credit Information Suppliers (ACCIS), https://www.ceps.eu/ceps-publications/european-credit-information-landscape/ (2011)

[20] Symanovich, S.: What is a credit freeze? should i freeze my credit? https://www.lifelock.com/learn-credit-finance-credit-freeze.html

[21] The European Consumer Organisation / Bureau Européen des Unions de Consommateurs (BEUC): The never-ending European credit data mess. https://www.beuc.eu/publications/beuc-x-2017-111_the-never-ending-european-credit-data-mess.pdf (2018)

[22] UK Parliament: Consumer Credit Act 1974, Part X: Ancillary credit businesses. https://www.legislation.gov.uk/ukpga/1974/39/part/X (2018)

[23] Vemprala, N., Dietrich, G.: A social network analysis (SNA) study on data breach concerns over social media. In: Proceedings of 52nd Hawaii International Conference on System Sciences (2019)

[24] Yang, S., Quan-Haase, A., Rannenberg, K.: The changing public sphere on Twitter: Network structure, elites and topics of the #righttobeforgotten. New Media & society **19**(12), 1983–2002 (2017)