

# Survey Report:

## Unveiling and Quantifying Facebook Exploitation of Sensitive Personal Data for Advertising Purposes

**Authors:** José González Cabañas, Ángel Cuevas, and Rubén Cuevas

**Publication:** Proceedings of the 27th USENIX Security Symposium

Project Report for Foundations of Privacy Course CS6413  
Faculty of Computer Science, University of New Brunswick,  
Fredericton, New Brunswick, Canada

Report By:

Kingsley Larbi (Student Number: 3574875)

Tolulope Olugbenga (Student Number: 3643581)

Kwasi Boakye-Boateng (Student Number: 3632590)

**Abstract**—The General Data Protection Regulation(GDPR) is a law whose main aim is to strengthen and unify personal data protection for all individuals living in the European Union(EU). GDPR came into full force in May 2018 and it restricts the processing of certain categories of personal data referred to as sensitive personal data. This data ranges from health, political orientation, sexual preferences, religious beliefs to ethnic origin and the likes. The authors of the title, sought to investigate whether Facebook were exploiting sensitive data of citizens for commercial purposes. It also sought to quantify the number of citizens that were affected before the implementation of GDPR. The authors placed emphasis on online advertising because it is seen by the them as the most important source of revenue for these companies. The authors noted that 73% of EU citizens were affected. We observed, in Canada, that investigations have taken place but Facebook has not been found culpable. We also observed that no research exists in investigating whether Canadian privacy laws such as PIPEDA have been violated by Facebook.

**Index Terms**—privacy, security, sensitive information.

### I. INTRODUCTION

General Data Protection Regulation (GDPR) [1] which came into effect in May 2018 due to European Union’s (EU) citizens’ concern about the handling of their personal information by online services/businesses. The author cites statistics by Special Eurobarometer 431 [2] about data protection which shows that 50% of EU citizens do not like provision of free services in exchange for their personal information, a little over three-fifths of the citizens do not trust online businesses and more than half of the citizens do not like tailored advertising based on their personal information from online businesses.

There are some categories of personal data tagged by GDPR as well as some previous EU national data protection laws as sensitive and they were tagged as “Specially Protected

Data”. “Special Categories of Personal Data” or “Sensitive Data”. These categories are prohibited by these laws with some exceptions such as the user has to provide explicit consent before the data was used for their intended purpose. GDPR defines sensitive data as follows “data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person’s sex life or sexual orientation”.

The paper sought to investigate whether online companies like Facebook [3] were exploiting sensitive data of citizens for commercial purposes. It also sought to quantify the number of citizens that were affected before the implementation of GDPR. The paper placed emphasis on online advertising because it is seen by the authors as the most important source of revenue for the companies. The authors also made a notable mention that Facebook was beaten only by Google in terms of advertising revenue [4].

### II. DETAILED DESCRIPTION OF THE TOPIC

In this section a brief background on Facebook and its advertisement policies, a brief background GDPR, the methodology and results of investigations into whether Facebook would have violated GDPR, had it been implemented before May 2018, are discussed.

#### A. Background

1) *General Data Protection Regulation(GDPR)*: General Data Protection Regulation(GDPR) which regulates all EU-member countries came into effect in May 2018. The main purpose of Article 9 of GDPR is about Sensitive Personal Data and its use. Article 9, which is entitled “*Processing of*

*Special Categories of Personal Data*”, states that “*Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs or trade union membership and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person’s sex life or sexual orientation shall be prohibited.*” There are some exceptions provided by GDPR, such as explicit user consent, before such a prohibition can be waived.

2) *Facebook Ad Preferences*: Facebook assigns ad preferences to users based on their activity on Facebook, websites and other online platforms. A user can edit his or her ad preferences at any time, but unfortunately only a few users know about this option. Facebook provided the following reasons why ad preferences are assigned to users:

- Ad preference was added by the user
- Ad preference is related to the user’s general activity on Facebook
- Ad preference is related to an ad clicked by the user in the past
- Ad preference is related to an app installed by the user
- Ad preference is related to a page liked by the user
- Ad preference is related to the comments, posts, shares or reactions of the user

Facebook labels users with so-called ad preferences based on their users’ online activity on the Facebook platform that Facebook tracks. So based on certain activities, users are tagged with certain preferences which are potentially inferred as sensitive information on the users. The authors states an example of an advertisement that was received by one of them which connoted that the person was a homosexual even though the author had not explicitly stated any sexual preference.

Clearly this assignment of potentially sensitive-information-based preferences is a violation of the GDPR since Facebook did not acquire explicit consent from the user before assigning such preferences. Such activities attracted a fine of 1.2M from the Spanish Data Protection Agency (DPA) after arguing that Facebook “collects, stores and uses data, including specially protected data, for advertising purposes without obtaining consent.” [5]

3) *Facebook Ads Manager*: The Facebook Ads Manager [6] is a tool that allows advertisers to reach their potential customers. Advertisers are given a wide range of configuration parameters in which they could specify the type of customers they would love to reach, such parameters include location, age, gender, behaviours, interests and so on. The interest parameter was the main focus of the paper, to help identify potential sensitive interests associated with Facebook users. Advertisers define their target customers based on these parameters, for example a makeup company could want to reach women living in Vancouver with interests in beauty/skin care. The Facebook Ads manager provides a potential reach of people matching the defined audience.

4) *Facebook Data Valuation Tool (FDVT)*: FDVT [7] is a web browser extension currently available to users on Google Chrome and Mozilla Firefox. It provides users with real

time data about how much revenue they have generated for Facebook due to their profile, the number of ads they click and see. As of Feb 2018 more than 6,000 users have installed the extension. FDVT collects the ad preferences assigned to a user by Facebook, this information could be used to find potential sensitive ad preferences that has been assigned to the FDVT user.

## B. Research Objectives

The authors’ research objective was to examine Facebook’s use of potentially sensitive data from January 2018 to May 2018, quantify the number of EU-based citizens (and their estimated percentage of the entire EU citizenry) that were assigned these potentially sensitive-information-based preferences. The authors created a web browser extension that collected data from Facebook users with their permission after installing it in their web browsers. The extension, known as the Data Valuation Tool for Facebook Users (FDVT) collects (among other data) the ad preferences Facebook assigns to users. Using the tool, the authors analysed more than 5.5 million ad preferences assigned to more than 4500 Facebook users. Out of the many ad-preferences assigned, approximately 126,000 were identified as unique. This process was undertaken to ensure that the preferences were assigned to real users.

The paper is grouped into three stages. In the first stage, the authors used natural different techniques identify potential sensitive-information-based preferences from the dataset acquired. Second, the preferences were used to query Facebooks’ Ads Manager to identify the number of users in the EU. Third, the authors extended the functionality of FDVT to warn Facebook users of assignments of such preferences. In summary, the authors identified the following:

- 2092 potentially sensitive ad preferences from the about 126,000 preferences.
- Facebook assigns an average of 16 potentially sensitive ad preferences to FDVT users
- More than 73% of EU-based Facebook users are labelled with at least one of the most popular sensitive ad preferences. These users make up 40% of EU citizens.
- Women have a significantly higher exposure than men to potentially sensitive ad preferences.
- People within the age group of 20-39 years old have the highest exposure of any age group.
- A ball-park estimation that suggests that unveiling the identity of Facebook users labeled with potentially sensitive ad preferences may be as cheap as 0.015 per user.

## C. Methodology

1) *Obtaining The Dataset*: The authors sought to collect a dataset of ad preferences linked to actual EU-based Facebook accounts to uncover potentially sensitive ad preferences and identify which of the accounts are assigned these preferences. The authors used FDVT to collect add preferences from users that had install it. They obtained a dataset from users who had installed FDVT between October 2016 and October 2017. The dataset includes ad preferences from a total of 4577 users out

of which 3166 are EU citizens. The total users were assigned 5.5M ad preferences out of which 126,192 are unique.

The dataset had the following information:

- *ID of the ad preference*: This key was used to identify the ad preference regardless of the language due to the ad preference being specified by users in different languages.
- *Name of the ad preference*: It was the main descriptor of the ad preference and was translated in English because Facebook returned names in English.
- *Disambiguation Category*: It was used to distinguish ad preferences that had the same name but different meanings. For example ,Nouveau (drink), Nouveau (clothing). The authors identified 700 of them.
- *Topic Category*: Some topics were assigned to ad preferences. For example, Juventus was linked to Sports.
- *Audience Size*: This value reports the number of Facebook users that have been assigned the ad preference worldwide.
- *Reason why the ad preference is added to the user*: The reason why the ad preference has been assigned to the user according to Facebook. There are six possible reasons

Each FDVT user is assigned a median of 474 preferences. For instance, each ad preference is assigned to a median of only 3 (0.06%) FDVT users. However, it is important to note that many ad preferences still reach a reasonable portion of users. Our dataset includes 1000 ad preferences that reach at least 11% of FDVT users.

## 2) Identification of Potentially Sensitive Ad Preferences:

In a bid to quantify the number of EU Facebook users that have been assigned potentially sensitive ad preferences, two processes were implemented to identify potentially sensitive ad preferences out of the 126K unique ad preferences assigned to FDVT users. In the first step, due to the large size of the dataset, Natural Language Processing techniques are combined with manual classification to obtain a list of likely sensitive ad preferences. In the second step, the list is fed into Facebook Ads Manager API to produce the number of Facebook users in each EU country that have been assigned at least one of the potentially ad preferences.

*a) Step 1A: Natural Language Processing*: To automatically identify likely sensitive ad preferences, the authors selected five of the relevant categories listed as Sensitive Personal Data by the GDPR:

- Data revealing racial or ethnic origin.
- Data revealing political opponents.
- Data revealing religious or philosophical beliefs.
- Data concerning health.
- Data concerning sex life and sexual orientation.

The authors selected those categories because of the likelihood of ad preferences revealing information related to them after they performed a preliminary manual inspection. They indicated that ad preferences they observed, such as "Socialism" and "Islam" fo may suggest political opinions and religious belief of Facebook users. To automatically identify

likely sensitive ad preferences, the authors sought perform a semantic mapping between ad preference and one of the five sensitive categories mentioned above. They achieved this by defining a dictionary including keywords and sentences that represented each of the five categories. Two sources of data were used to create this dictionary. The first source was a list of controversial issues obtained in Wikipedia [8] and the second was from the Datamuse API [9]. Politics and economics, religion, and sexuality categories were selected. Using Data Muse API, the authors fetched words that were semantically similar in meaning to the five categories. The combination of the two sources resulted in the dictionary having a total of 264 keywords which were used in a semantic similarity computation to identify, out of the 126K ad preferences, ones that were likely to be sensitive. Each ad preferences was compared with the 264 words in the dictionary and the highest similarity score out of the 264 was recorded for that ad preference. Spacy [10] was used for this computation with the similarity value ranging between 0 and 1, with 1 being the highest and 0 the lowest. It was observed that even though an ad preference closely matched a dictionary word, the similarity score was very low. For example, names of politicians which could reveal a preference to politics such as *<name: "Angela Merkel" , disambiguation: Politician>* had low scores. To rectify this, ad preferences that had a disambiguation field, had that field used instead of the actual name of the preference. Fig 1 shows the cumulative distribution function (CDF) for the semantic similarity scores of the 126K ad preferences. The authors chose a threshold of 0.6 due to the trend of the CDF and that resulted in the number of ad preferences being scaled down from 126K to 4452.

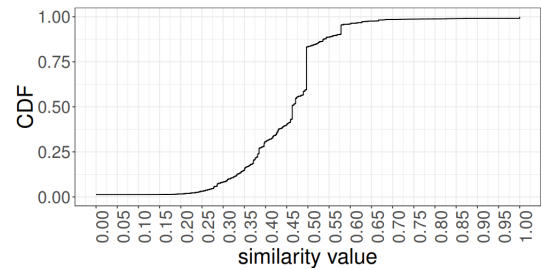


Fig. 1. Cumulative Distribution Function for Semantic Similarity Scores Assigned to 126K Ad Preferences [11]

*b) Step 1B: Manual Classification*: Twelve panelists, knowledgeable in the area of privacy, were selected to each manually classify a random sample (between 1000 and 4452 elements) from the 4452 ad preferences. They were requested to classify each ad preference into one of the five categories or a "Not known" category, if the panelist does not know the meaning of the ad preference. To aid in the manual classification, each researcher was given all contextual information Facebook offers per ad preference: name, disambiguation category (if available) and topic (if available). Each ad-preference was classified by 5 panelists and majority voting (3 or more) was used to classify each ad preferences as either

TABLE I  
AD PREFERENCES PER NUMBER OF VOTES [11]

Number of Votes	Ad Preferences
0	1054
1	767
2	539
3	422
4	449
5	1221

sensitive or non-sensitive. Based on the results in Table I, 2092 out of the 4452 ad preferences were labelled as sensitive which represented 1.66% of the 126K dataset. The Fleiss' Kappa test [12] [13] was used by the authors to evaluate the voters' agreement and a Fleiss' Kappa coefficient of 0.94 was obtained implying that the panelists were in an almost perfect agreement. The distribution of the 2092 ad preferences across the five sensitive categories were as follows: 58.3% are related to politics, 20.8% to religion, 18.2% to health, 1.5% to sexuality, 1.1% to ethnicity and just 0.2% present discrepancy among votes.

c) *Step 2: Retrieving Number of Facebook Users Assigned With Potentially Sensitive Ad Preferences:* In January 2018 (before GDPR came into effect), the authors used Facebook Ads Manager API to retrieve the number of Facebook users, that have been tagged with each of the 2092 potentially sensitive ad preferences, in each EU country. From the results, the ad preferences were sorted in decreasing order of popularity, in each country, to allow the authors to compute the number of Facebook users tagged with per Top  $N$  ( $N$  ranges between 1 and 2092) potentially sensitive ad preferences. For example, using the OR feature in Facebook Ads Manager API and  $N = 3$ , query like “how many people in Greece are interested in Christianity OR Islam OR Veganism” can be made. Because the count of users would not be a fair assessment to all countries due to variations in size, the users rather used percentages. They calculated that using the following metrics:

- $FFB(C, N)$ : Percentage of users in a country (or whole EU),  $C$ , that have been tagged with least one of the top  $N$  potentially sensitive ad preferences. It is calculated as the ratio between the users that have been tagged in  $C$  and the total number of users in  $C$ . Total number of users in  $C$  can be retrieved from the Facebook Ads Manager API.
- $FC(C, N)$ : Percentage of citizens in  $C$ , that have been tagged with at least one of the top  $N$  potentially sensitive ad preferences. It is computed as the ratio between the number of citizens that been tagged in  $C$  and the total population of  $C$ . The World Bank data was used to EU countries' populations.

The authors used popularity as a criterion to select the top  $N$  ad preferences out of the 2092 potentially sensitive ad preferences. They selected the  $N$  ad preferences assigned to the most users according to the Facebook Ads Manager API. They gave two reason why there is a likelihood of

TABLE II  
FREQUENCY OF REASONS AD PREFERENCES WERE ASSIGNED BY FACEBOOK [11]

Reason Of Assignment	All Ad Preferences	Potentially Sensitive Ones
Due to a like	71.64%	81.36%
Due to an ad click	21.51%	15.85%
FB suggests it could be relevant	4.83%	2.45%
Due to an app installation	1.78%	0.04%
Due to comments or reaction buttons	0.18%	0.26%
Added by user	0.04%	0.03%
Unclear or not gathered by FDVT	0.01%	0.01%

$FFB(C, N)$  and  $FC(C, N)$  reporting a lower bound concerning the total percentage of Facebook users and citizens in  $C$ , tagged with potentially sensitive ad preferences. First, these metrics use at most  $N = 2092$  potentially sensitive ad preferences, which (assuming that the voters were accurate) is very likely a subset of all sensitive ad preferences available on Facebook. Second, the Facebook Ads Manager API only allowed creating audiences with at most  $N = 1000$  interests (i.e., ad preferences). Beyond  $N = 1000$  interests, the API provided a fixed number of Facebook users independently of the defined audience. This fixed number was 2.1 billion which to the best of the authors' knowledge referred the total number of Facebook users included in the Ads Manager. Therefore, in practice, the maximum value of  $N$  that was chosen to use in FFB and FC was 1000.

#### D. Quantifying The EU Users

1) *FDVT Users:* According to the authors, 90% of FDVT users (which amounted to 4121) were tagged with at least one sensitive ad preference with the overall estimation of 2092 unique sensitive ad preferences being assigned to FDVT users more than 146K times. 2848 out of the FDVT users, tagged with potentially sensitive ad preferences, were EU users and they were assigned more than 100K sensitive interests with 1528 being unique. Table II shows the frequencies of reasons why, according to Facebook, each ad preference (including potentially sensitive ones) has been assigned to a user. According to the table, most of the sensitive ad preferences were derived from users likes (81%) or clicks on ads (16%). There were very few cases (0.03%) in which users proactively include potentially sensitive ad preferences in their list of ad preferences using the configuration setting offered by Facebook. This clearly shows that Facebook does not request for permission from users before potentially sensitive ad preferences were assigned.

2) *EU Facebook users and citizens:* According to Fig. 2, which shows the  $FFB(C, N)$  for values of  $N$  ranging between 1 and 1000, there was a stable value of users for  $N$  ranging between 500 and 1000, which was stable each EU country. This indicated that any user tagged with potentially sensitive ad preferences outside the top 500 was likely to

be already tagged with at least one potentially sensitive ad preference within the top 500. Thus  $N = 500$  was chosen. Table III shows  $FFB(C, N = 500)$  and  $FC(C, N = 500)$  for every EU country. The last row in the table shows average results for the 28 EU countries together. It was observed that 73% of EU Facebook users, which corresponds to 40% of EU citizens, are tagged with some of the top 500 potentially sensitive ad preferences. Focusing on individual countries,  $FC(C, N = 500)$  revealed that in 7 of them more than half of their citizens are tagged with at least one of the top 500 potentially sensitive ad preferences with Malta (66.37%) being the highest. Germany was the least impacted with 30.24%. Moreover,  $FFB(C, N = 500)$  has France being the lowest with 65% and Portugal being the highest with 81%. This implied that about two-thirds or more of Facebook users in any EU country are tagged with some of the top 500 potentially sensitive ad preferences. The results suggested that a very significant part of the EU population can be targeted by advertising campaigns based on potentially sensitive personal data.

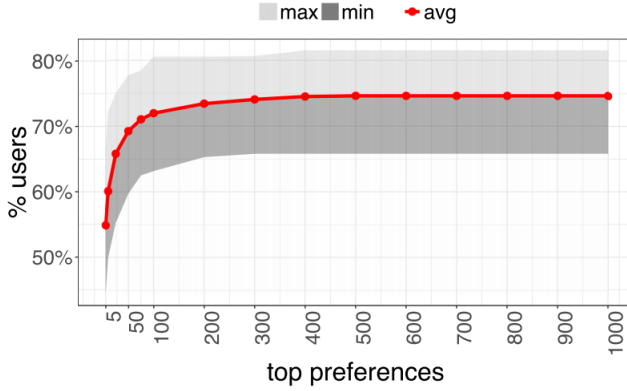


Fig. 2.  $FFB(C, N)$  for values of  $N$  between 1 and 100 [11]

3) *Expert-verified sensitive ad preferences*: In order to confirm that the set of potentially sensitive ad preferences contained ones likely relevant under GDPR, the authors examined a subset of 20 ad preferences that all panelists classified as sensitive and had an expert from the Spanish DPA review and confirm the sensitivity of each of the 20 ad preferences in that subset according to the GDPR, though the subset is not necessarily representative of all potentially sensitive ad preferences. Tables IV and V show the percentage of Facebook users ( $FFB(C, N = 20)$ ) and citizens ( $FC(C, N = 20)$ ) per EU country. The last row is the aggregate results for the 20 in each country, and the last column is the aggregate results for the 28 EU countries together. It was observed that 42.9% of EU Facebook users, which corresponds to 23.5% of EU citizens, are tagged with at least one of the expert-verified sensitive ad preferences. Hence, around one-quarter of the EU population has been tagged in Facebook with at least one of the expert-verified sensitive ad preferences. If the results were analyzed per country, it can be observed that the fraction of the population affected ranges between 15% in Estonia (EE),

TABLE III  
PERCENTAGE OF EU FACEBOOK USERS AND EU CITIZENS TAGGED [11]

country	C	FFB(C,500)	FC(C,500)
Austria	AT	75	37.73
Belgium	BE	70.27	45.82
Bulgaria	BG	72.97	37.88
Croatia	HR	80	38.36
Cyprus	CY	79.17	64.95
Czech Republic	CZ	71.7	35.98
Denmark	DK	77.5	54.09
Estonia	EE	66.67	36.46
Finland	FI	70.97	40.04
France	FR	65.79	37.37
Germany	DE	67.57	30.24
Great Britain	GB	75	50.28
Greece	GR	77.19	40.94
Hungary	HU	75.44	43.8
Ireland	IE	80.65	52.38
Italy	IT	79.41	44.55
Latvia	LV	72.53	33.67
Lithuania	LT	75	41.78
Luxemborg	LU	72.22	44.6
Malta	MT	80.56	66.37
Netherlands	NL	74.55	48.18
Poland	PL	75	31.62
Portugal	PT	81.54	51.33
Romania	RO	75.76	38.06
Spain	ES	74.07	43.06
Slovakia	SK	70.37	35
Slovenia	SI	78	37.78
Sweden	SE	73.97	54.53
<b>European Union</b>	<b>EU</b>	<b>73.25</b>	<b>40.63</b>

Latvia (LV) and Poland (PL) and 38% in Malta (MT). These findings suggest that Facebook may have used GDPR-relevant data for a large percentage of EU citizens in the period prior to when the GDPR became enforceable.

4) *Age and gender analysis*: The authors analyzed the association of different demographic groups (based on gender and age) with potentially sensitive ad preferences. The gender analysis considered two groups, men vs. women, while the age analysis considered four age groups namely Adolescence (13-19), Early Adulthood (20-39), Adulthood (40-64) and Maturity (65+), based on a division proposed by Erikson *et al.* [14]. For each group,  $FFB(C = EU28, N = 500)$  and  $FC(C = EU28, N = 20)$  were computed. Fig. 3 and Fig. 4 report the results for age and gender groups, respectively showing that the Early Adulthood group is clearly the most exposed age group. 61% (45%) of users in this group were tagged with some of the Top 500-suspected (20-expert-verified) sensitive ad preferences. In particular, more than one-quarter of the users within every group were exposed to expert-verified sensitive ad preferences. The gender-based analysis showed that 78% (49%) of women were exposed to the Top 500-suspected (20-expertverified) ad preferences which suggested the existence of a gender bias.

#### E. Commercial Exploitation of Sensitive Ad Preferences with Real FB Ad Campaigns

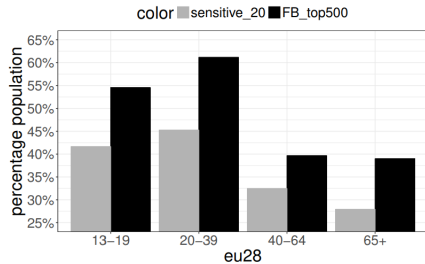
The authors wanted to show that a significant number of EU citizens were assigned some potentially sensitive ad

TABLE IV  
PERCENTAGE OF FFB PER EU COUNTRY [11]

Name	Communism	Islam	Quran	Suicide Prevention	Socialism	Judaism	Homosexuality	Alternative Medicine	Christianity	Illegal Immigration	Oncology	Light Community	Gender Identity	Reproductive Health	Bible	Pregnancy	Nationalism	Veganism	Buddhism	Feminism	Union
AT	0.48	8.18	3.41	0.14	1	2.5	6.14	5	10.68	0.17	0.23	6.36	0.03	0.01	17.95	15.68	0.86	14.55	3.18	4.55	45.45
BF	0.61	7.16	3.38	0.15	0.78	1.16	5.54	2.97	7.43	0.07	0.27	6.62	0.08	0.07	10.81	12.97	0.78	10.27	3.38	3.78	39.19
BG	1.35	4.59	1.08	0.2	0.57	0.86	2.97	8.38	6.22	0.1	0.62	5.14	0.01	0.2	8.65	9.19	1.65	7.3	1.62	3.51	32.43
HR	1.3	5.5	1	0.32	0.48	0.7	6.5	6	7.5	0.02	0.44	6.5	0.08	0.4	10.5	17	1.85	10.5	3.55	3.8	41.5
CY	1.67	13.54	4.48	0.21	1.15	2.29	4.38	5.62	9.69	0.07	3.96	6.56	0.88	0.02	11.46	13.54	2.19	10.21	3.33	5.52	45.83
CZ	3.21	4.91	0.45	0.12	2.45	0.72	5.47	4.15	3.77	0.68	0.57	6.04	0.02	0.14	7.17	16.23	2.45	9.25	2.26	2.08	37.74
DK	0.38	6.75	1.9	0.12	3	2.17	5	4	15	0.05	0.15	6.5	0.03	0.05	12.75	14.5	1	12.75	2.08	5.5	45
EE	0.61	2.22	0.65	0.1	0.76	1.01	3.89	4.17	2.22	0.01	0.1	5.14	0.02	0.02	4.31	10	0.58	9.86	1.53	2.78	27.78
FI	0.52	4.19	1.16	0.09	0.48	0.61	5.16	4.19	4.19	0.07	0.08	6.45	0.02	0.06	4.84	11.29	0.45	15.16	1.13	6.77	35.48
FR	2.29	7.89	3.95	0.16	0.47	1.26	7.37	2.89	5.53	0.05	0.17	7.11	0.07	0.01	7.63	10.79	1.08	8.68	2.61	5	34.21
DE	0.43	7.57	3.24	0.14	0.43	1.38	5.68	3.24	6.49	0.06	0.16	5.95	0.03	0.01	15.41	11.89	1	11.35	1.43	3.78	40.54
GR	0.81	4.21	1.18	0.23	0.91	1.3	5.09	7.19	6.67	0.26	0.49	5.79	0.56	0.04	8.25	13.51	1.74	9.82	2.63	3.68	36.84
HU	0.74	2.28	0.74	0.12	1.93	1.16	4.21	4.21	9.3	0.26	0.3	4.39	0.07	0.1	10	11.23	2.11	9.82	3.33	2.46	36.84
IE	0.52	4.19	1.35	1.1	1.1	1.26	9.03	9.68	10.97	0.06	1.29	11.94	0.23	0.71	19.03	20.97	2	14.84	3.87	9.35	51.61
IT	1.15	4.12	1.71	0.28	3.53	2.29	7.65	6.18	12.65	0.08	0.94	8.53	0.07	0.04	17.65	12.35	1.32	13.53	2.94	5.88	44.12
LV	0.56	2.75	1.01	0.13	0.34	1.76	4.62	3.96	3.19	0.02	0.7	5.27	0.2	0.07	5.71	13.19	2.42	9.23	1.98	3.19	32.97
LT	0.94	2.38	0.51	0.15	0.94	1.81	3.19	2.56	3.81	0.06	1.62	5.88	0.1	0.05	6.25	18.75	0.94	8.12	1.88	3.56	36.25
LU	0.64	5	1.83	0.28	2.78	1.19	5	5.56	7.22	0.01	0.19	6.67	0.1	0.01	14.44	12.78	2.19	13.06	3.33	5.83	41.67
MT	0.39	6.67	1.86	0.27	1.08	3.06	7.5	7.5	18.89	0.08	0.78	9.44	0.14	0.24	20.28	9.72	2.78	13.33	4.17	8.61	47.22
NL	0.24	5.36	2.45	0.15	0.28	1	6.18	3.64	5.18	0.02	0.45	6.36	0.03	0.03	10.91	14.55	0.7	10.91	2.45	3.64	40
PL	2.19	2.44	0.45	0.14	0.5	1.19	3.56	2.25	6.25	0.02	1.25	5.88	0.05	0.01	14.38	15	3	8.12	1.31	3.44	36.88
PT	0.94	3.69	0.62	0.22	2.15	1.69	4.46	8	12.46	0.02	1.09	7.85	0.05	0.04	12.31	18.46	1.69	11.23	6.92	8.15	44.62
RO	1.9	3.5	0.77	0.13	0.35	1.4	3.8	3.9	10	0.02	0.73	6.3	0.04	0.01	8.7	9.7	2.5	6.7	1.9	2.4	34.34
SK	1.74	3.11	0.56	0.44	2.33	0.93	4.44	2.93	4.81	0.11	0.59	4.81	0.01	0.03	6.67	18.89	1.37	8.52	1.67	4.07	35.56
SI	1.7	6.5	2	0.26	0.82	0.74	7.6	5	4.6	0.36	0.21	6	0.08	0	7.4	13	0.61	14	3	3.9	39
ES	0.56	4.07	0.96	0.44	1.48	1.15	8.15	5.56	10	0.14	0.7	7.04	0.07	0.03	7.04	14.07	1.11	10.37	2.19	8.89	40.74
SE	0.3	6.58	2.74	0.15	1.37	0.64	4.93	3.84	4.66	0.33	0.08	6.44	0.09	0.05	3.48	13.42	0.99	16.44	1.51	13.7	41.1
GB	0.41	6.82	3.64	0.27	0.93	0.95	8.64	6.14	7.5	0.05	0.66	11.14	0.55	0.13	15.68	18.41	0.91	13.64	2.5	7.27	47.73
EU28	0.93	5.71	2.46	0.28	1.21	1.32	6.79	4.29	8.21	0.09	0.61	8.21	0.1	0.07	12.14	14.29	1.39	11.43	2.39	7.5	42.86

TABLE V  
PERCENTAGE OF FC PER EU COUNTRY [11]

Name	Communism	Islam	Quran	Suicide Prevention	Socialism	Judaism	Homosexuality	Alternative Medicine	Christianity	Illegal Immigration	Oncology	Light Community	Gender Identity	Reproductive Health	Bible	Pregnancy	Nationalism	Veganism	Buddhism	Feminism	Union
AT	0.24	4.12	1.71	0.07	0.5	1.26	3.09	2.52	5.37	0.09	0.11	3.2	0.01	0	9.03	7.89	0.43	7.32	1.6	2.29	22.86
BF	0.4	4.67	2.2	0.1	0.51	0.76	3.61	1.94	4.85	0.04	0.18	4.32	0.05	0.05	7.05	8.46	0.51	6.7	2.2	2.47	25.55
BG	0.7	2.39	0.56	0.1	0.29	0.45	1.54	4.35	3.23	0.05	0.32	2.67	0.01	0.11	4.49	4.77	0.86	3.79	0.84	1.82	16.84
HR	0.62	2.64	0.48	0.15	0.23	0.34	3.12	2.88	3.6	0.01	0.21	3.12	0.04	0.19	5.04	8.15	0.89	5.04	1.7	1.82	19.9
CY	1.37	11.11	3.67	0.17	0.94	1.88	3.59	4.61	7.95	0.06	3.25	5.38	0.72	0.02	9.4	11.11	1.79	8.38	2.73	4.53	37.6
CZ	1.61	2.46	0.23	0.06	1.23	0.36	2.75	2.08	1.89	0.34	0.28	3.03	0.01	0.07	3.6	8.14	1.23	4.64	1.14	1.04	18.94
DK	0.26	4.71	1.33	0.08	2.09	1.52	3.49	2.79	10.47	0.03	0.1	4.54	0.02	0.04	8.9	10.12	0.7	8.9	1.45	3.84	31.41
EE	0.33	1.22	0.36	0.05	0.42	0.55	2.13	2.28	1.22	0	0.06	2.81	0.01	0.01	2.35	5.47	0.32	5.39	0.84	1.52	15.19
FI	0.29	2.37	0.66	0.05	0.27	0.35	2.91	2.37	2.37	0.04	0.05	3.64	0.01	0.04	2.73	6.37	0.25	8.55	0.64	3.82	20.02
FR	1.3	4.48	2.24	0.09	0.27	0.72	4.19	1.64	3.14	0.03	0.1	4.04	0.04	0.01	4.34	6.13	0.61	4.93	1.48	2.84	19.43
DE	0.19	3.39	1.45	0.06	0.19	0.62	2.54	1.45	2.9	0.03	0.07	2.66	0.01	0	6.9	5.32	0.45	5.08	0.64	1.69	18.14
GR	0.43	2.23	0.62	0.12	0.48	0.69	2.7	3.82	3.54	0.14	0.26	3.07	0.3	0.02	4.37	7.16	0.92	5.21	1.4	1.95	19.54
HU	0.43	1.32	0.43	0.07	1.12	0.67	2.44	2.44	5.4	0.15	0.17	2.55	0.04	0.06	5.81	6.52	1.22	5.7	1.94	1.43	21.39
IE	0.34	2.72	0.88	0.71	0.71	0.82	5.87	6.29	7.12	0.04	0.84	7.75	0.15	0.46	12.36	13.62	1.3	9.64	2.51	6.08	33.52
IT	0.64	2.31	0.96	0.16	1.98	1.29	4.29	3.47	7.1	0.04	0.53	4.79	0.04	0.02	9.9	6.93	0.74	7.59	1.65	3.3	24.75
LV	0.26	1.28	0.47	0.06	0.16	0.82	2.14	1.84	1.48	0.01	0.33	2.45	0.09	0.03	2.65	6.12	1.12	4.28	0.92	1.48	15.3
LT	0.52	1.32	0.28	0.08	0.52	1.01	1.78	1.43	2.12	0.03	0.91	3.27	0.06	0.03	3.48	10.44	0.52	4.53	1.04	1.98	20.19
LU	0.39	3.09	1.13	0.17	1.72	0.74	3.09	3.43	4.46	0.01	0.12	4.12	0.06	0.01	8.92	7.89	1.36	8.06	2.06	3.6	25.73
MT	0.32	5.49	1.53	0.22	0.89	2.52	6.18	6.18	15.56	0.07	0.64	7.78	0.12	0.19	16.71	8.01	2.29	10.99	3.43	7.09	38.91
NL	0.15	3.47	1.59	0.09	0.18	0.65	4	2.35	3.35	0.01	0.29	4.11	0.02	0.02	7.05	9.4	0.45	7.05	1.59	2.35	24.85
PL	0.92	1.03	0.19	0.06	0.21	0.5	1.5	0.95	2.64	0.01	0.53	2.48	0.02	0.01	6.06	6.32	1.26	3.43	0.55	1.45	15.55
PT	0.59	2.32	0.39	0.14	1.36	1.07	2.81	5.04	7.85	0.01	0.69	4.94	0.03	0.02	7.75	11.62	1.07	7.07	4.36	5.13	28.09
RO	0.96	1.78	0.39	0.07	0.18	0.71	1.93	1.98	5.07	0.01	0.37	3.2	0.02	0.01	4.42	4.92	1.27	3.4	0.96	1.22	17.25
SK	0.87	1.55	0.28	0.22	1.16	0.46	2.21	1.46	2.39	0.05	0.29	2.39	0	0.01	3.32	9.39	0.68	4.24	0.83	2.03	17.68
SI	0.82	3.15	0.97	0.13	0.4	0.36	3.68	2.42	2.23	0.17	0.1	2.91	0.04	0	3.58	6.3	0.3	6.78	1.45	1.89	18.89
ES	0.32	2.37	0.56	0.26	0.86	0.67	4.74	3.23	5.81	0.08	0.41	4.09	0.04	0.02	4.09	8.18	0.65	6.03	1.27	5.17	23.68
SE	0.22	4.85	2.02	0.11	1.01	0.47	3.64	2.83	3.43	0.24	0.06	4.75	0.06	0.03	4.04	9.9	0.73	12.12	1.11	10.1	30.29
GB	0.27	4.57	2.44	0.18	0.62	0.64	5.79	4.11	5.03	0.04	0.44	7.47	0.44	0.09	10.51	12.34	0.61	9.14	1.68	4.88	31.99
EU28	0.51	3.13	1.35	0.15	0.66	0.72	3.71	2.34	4.49	0.05	0.33	4.49	0.33	0.04	6.64	7.82	0.76	6.25	1.31	4.1	23.45





beliefs, political opinions, sexual orientation, between October 6th and October 15th, 2017; these ads were focused on France, Germany, Italy and Spain. The authors spent 35 euros on the ad campaigns, with it they were able to reach 26458 tagged users. They authors concluded; the experiment proves that FB generated revenue from its users while exploiting their sensitive personal data for commercial purposes.

#### F. Ethics And Privacy Risks Associated With Sensitive Personal Data Exploitation

Facebook makes it very easy to reach people from different places with various interests. This makes it easy for an adversary to target a certain group of FB users. Examples include:

- *Hate campaigns*: An adversary could create a hate speech campaign to target people with specific sensitive interests. E.g. Offensive messages to users tagged with homosexuality.
- *Identification attack*: An attacker could create an ad campaign to identify people in a sensitive social group, such as religious belief, sexual orientation, etc. The attacker could create an ad campaign about a survey that if completed puts them in a raffle to win an iPhone X. Once the user is on the attacker's website, the attacker could employ different procedures to collect the user's personal information that reveals his/her identity [15].

A study based on email phishing showed that 9% of users put their login details on phishing websites [16]. The authors spent 35 euros and reached 26k users; based on the 9% success rate, the authors would identify 2.34k users if they wanted to. This discovery reveals that it costs as low as 0.015 euro to identify a user in a sensitive social group from FB ads. This cost is so low as compared to the security risks it poses to users. For example, (i) In a country where Homosexuality is illegal, the government could identify people likely homosexual in such country; (ii) Health insurances companies could identify people with bad habits, health problems and reject or charge them higher even if they may have been labelled wrongly.

#### G. FDVT Extension To Inform Users About Their Potentially Sensitive Ad Preferences

The findings in the previous section motivated the authors to create FDVT browser extension that informs users about the potentially sensitive ad preferences that they have been labelled with. They built a classifier that classifies ad preferences as sensitive or non-sensitive. During the installation process users have to read and accept the Terms of Use [17] and privacy policy, users also have to grant explicit permission for the information collected to be used for research purposes.

*1) Automatic Binary Classifier For Sensitive Ad Preferences*: In order to build an automatic binary classifier, the authors defined a threshold where ad preferences above or below were classified as sensitive and non-sensitive respectively. The authors used the automatically filtered dataset of 4452 ad preferences. 80% of the dataset was used in training and the remaining 20% was used for validation/testing. They

validated every selected threshold by computing the precision, recall and F-score on the validation subset. Different validation subsets were used to prove the firmness of the binary classifier, 5,000 iterations were made. Fig. 5 shows the AUC, precision, recall and F-score after 5,000 iterations. The classifier was quite stable when the optimal threshold was in range 0.68 - 0.69, the AUC shows a really stable result around 0.86. After testing, it was found that the median precision of the classifier was 0.837 with a min and max of 0.75 and 0.90 respectively; the median recall was 0.78 with a min and max of 0.70 and 0.86 respectively.

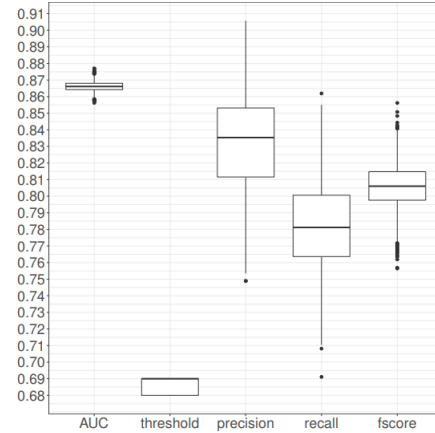


Fig. 5. AUC, precision, recall and F-score for the optimal threshold to automatically classify an ad preference as sensitive or non-sensitive [11]

#### 2) System Implementation:

*a) FDVT Backend*: Ad preferences whose similarity score was greater than or equal to 0.69 were classified as sensitive and added to a blacklist. Each time a user logs in to FB the system gets an updated set of as preferences, it compares this set with what is contained in the blacklist. The system stores the list of potentially ad preferences assigned to the user and notifies them when FB has removed it. Anytime a new ad preference is assigned that's not in the FDVT database, the similarity score is computed and included in the blacklist if it's classified as sensitive.

*b) FDVT User Interface*: FDVT has a button that shows a list of potentially sensitive ad preferences assigned to a user, an example could be seen in Fig 6. This list includes Ad preference name, Date of Addition, Date of Deletion (If deleted), Description of why it was assigned, Status (Active or Deleted)

Potentially sensitive interests in your profile:				
Preference Name	Addition	Deletion	Description	Status
Democracy	2017-06-12	--	You have this preference because you liked a Page related to Democracy.	Active
Homosexuality	2017-09-25	--	You have this preference because you liked a Page related to Homosexuality.	Active
Socialism	2017-09-28	--	You have this preference because you liked a Page related to Socialism.	Active
Veganism	2017-11-18	--	You have this preference because you clicked a Page related to Veganism.	Active
Bible	2017-12-23	--	This is a preference you added.	Active
Pregnancy	2017-05-20	2017-07-10	You have this preference because you installed an app related to Pregnancy.	Deleted
Quran	2017-05-20	2017-08-30	You have this preference because you liked a Page related to Quran.	Deleted

Fig. 6. Webpage displaying sensitive preferences [11]

## H. Authors' Conclusion

These findings by the authors reveal that Facebook exploits potentially sensitive information of its users for commercial purposes. Facebook was fined in France and Spain for this practice; recently Facebook was fined in Italy for misleading users to sign up without explicitly informing them that their information would be sold to third parties [18]. After studying the potentially sensitive personal data assigned by FB to EU users before the enforcement of GDPR (May 25, 2018), the results show that 73% of EU FB users corresponding to 40% of EU citizens were affected. The authors observed the risks assigned ad preference posed to FB users like hate speech attack or identification attacks. They concluded that, Facebook should desist for labelling its users with all ad preferences that can be used to deduce a user's religious beliefs, ethnic origin, sexual orientation, political orientation, health conditions. This would protect Facebook for unnecessary fines due to the violation of GDPR and it would protect its users from privacy risks associated with the exploitation of their sensitive data.

## III. RELATED WORK IN THIS AREA

Carrascosa *et al* [19] investigated how prospects are targeted and labelled based on their online activities for advertising purposes. They proposed a new method to appraise the amount of ads targeted to users when they surf the internet. Bots were created to quantum the number of received ads which actually matches the specific interest of a user. Their analysis after checking over 50 users indicates that the volume of behavioural targeted ads received by a user depends on the economic value associated to the interests of the user. Their findings reveals that advertising companies targets behavioural traits related to sensitive topics irrespective of existing acts against this.(e.g In Europe). Castelluccia *et al* [20] categorized and filtered targeted ads, they showed that targeted ads contains information that could be used to accurately create an interest profile for users. They argued that with a limited number of targeted ads, an adversary could deduce user interests by creating an interest profile that reveals up to 58% of a user actual interests with an accuracy of more than 79%. The authors believe that if these unveiled interests are sensitive, it could led to serious privacy issues.

## IV. STUDENTS' THOUGHTS AND REMARKS

### A. Facebook and Privacy-Related Issues in Canada

It was observed that there was no research about in such a similar manner pertaining to Canadian citizens. We read news reports in which Facebook reached an agreement with the Office of the Privacy Commissioner of Canada not to use non-users' personal information in users' address books [21]. There was another publication in which the Privacy Commissioner of Canada launched an investigation into alleged unauthorized access and use of Facebook user profiles in the wake of the Cambridge Analytica scandal [22]. In all these, most governments should take a strong stand in ensuring the privacy of users especially with regards to sensitive data.

### B. FDVT - Who Watches The Watchman?

Is the data collected from FDVT users safe in the hands of the authors? We know during the installation process users have to agree that their data collected will be used for research purposes. We do not know whether they will be audited on consistent basis to ensure that they stay true to their word. Then again, Facebook has auditing policies however the paper by the authors have confirmed issues with regards to privacy. If Facebook cannot be trusted, what about the watchman (FDVT)?

### C. Trust and Social/Online Companies.

Many online companies use our personal data for online advertising in exchange of giving us free services to use their platform. We use most of these online services for our day to day life. It is very difficult to trust these companies so as users we must be very circumspect in how we use their products and not expose everything online.

### D. The sensitive data collected, are they accurate?

And lastly, we ran the FDVT classifier across various accounts. We noticed a certain trend, some of the potentially sensitive ad preferences assigned to users were not accurate. Given the serious implications and privacy risks an ad preference can pose a user, it's even worse that users may have to suffer due to wrongfully assigned ad preferences. For instance, a user may be tagged with "Drinking in the Night (Wrongfully)", "Driving", "Cinemas"; If this user goes to the cinema every Friday night, it could be concluded that the user drinks and drives as well. If there is an accident during this period and the police has access to this information, the user could potentially become a suspect of a murder.

## REFERENCES

- [1] "L\_20161119EN.01000101.xml." [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN>
- [2] "Special Eurobarometer 431 DATA PROTECTION REPORT Special Eurobarometer 431 / Wave EB83.1-TNS opinion & social." [Online]. Available: [http://ec.europa.eu/public\\_opinion/index\\_en.htm](http://ec.europa.eu/public_opinion/index_en.htm)
- [3] "Facebook - Log In or Sign Up." [Online]. Available: <https://www.facebook.com/>
- [4] "Google and Facebook Tighten Grip on US Digital Ad Market - eMarketer." [Online]. Available: <https://www.emarketer.com/Article/Google-Facebook-Tighten-Grip-on-US-Digital-Ad-Market/1016494>
- [5] "Facebook fined 1.2M for privacy violations in Spain — TechCrunch." [Online]. Available: <https://techcrunch.com/2017/09/11/facebook-fined-e1-2m-for-privacy-violations-in-spain/>
- [6] "Log into Facebook — Facebook." [Online]. Available: <https://www.facebook.com/ads/manager>
- [7] J. González Cabañas, A. Cuevas, and R. Cuevas, "Fdvt: Data valuation tool for facebook users," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 2017, pp. 3799–3809.
- [8] "Wikipedia:List of controversial issues - Wikipedia." [Online]. Available: [https://en.wikipedia.org/wiki/Wikipedia:List\\_of\\_controversial\\_issues](https://en.wikipedia.org/wiki/Wikipedia:List_of_controversial_issues)
- [9] "Datamuse API." [Online]. Available: <https://www.datamuse.com/api/>
- [10] "spaCy · Industrial-strength Natural Language Processing in Python." [Online]. Available: <https://spacy.io/>
- [11] J. G. Cabañas, A. Cuevas, and R. Cuevas, *Unveiling and Quantifying Facebook Exploitation of Sensitive Personal Data for Advertising Purposes*. [Online]. Available: <https://www.usenix.org/conference/usenixsecurity18/presentation/cabanas>



- [12] J. L. Fleiss, B. Levin, and M. C. Paik, *Statistical methods for rates and proportions*. J. Wiley, 2003. [Online]. Available: <https://www.wiley.com/en-ca/Statistical+Methods+for+Rates+and+Proportions,+3rd+Edition-p-9780471526292>
- [13] J. L. Fleiss, "Measuring nominal scale agreement among many raters." *Psychological Bulletin*, vol. 76, no. 5, pp. 378–382, 1971. [Online]. Available: <http://content.apa.org/journals/bul/76/5/378>
- [14] E. H. Erikson and J. M. Erikson, *The life cycle completed (extended version)*. WW Norton & Company, 1998.
- [15] J. Hong, "The state of phishing attacks," *Commun. ACM*, vol. 55, no. 1, pp. 74–81, 2012.
- [16] X. Han, N. Kheir, and D. Balzarotti, "Phisheye: Live monitoring of sandboxed phishing kits," in *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 2016, pp. 1402–1413.
- [17] "FDVT - Terms of Use." [Online]. Available: [https://fdvt.org/terms\\_of\\_use/](https://fdvt.org/terms_of_use/)
- [18] "FACEBOOK FINED £8.9M BY ITALY FOR MISLEADING USERS OVER DATA USE." [Online]. Available: <https://www.independent.co.uk/life-style/gadgets-and-tech/facebook-fine-italy-user-data-scandal-privacy-settings-cambridge-analytica-social-media-a8673376.html>
- [19] J. M. Carrascosa, J. Mikians, R. Cuevas, V. E. Guavus, and N. Laoutaris, "I Always Feel Like Somebody's Watching Me Measuring Online Behavioural Advertising." [Online]. Available: <http://dx.doi.org/10.1145/2716281.2836098>
- [20] C. Castelluccia, M.-A. Kaafar, and M.-D. Tran, "Betrayed by Your Ads!" Springer, Berlin, Heidelberg, 2012, pp. 1–17. [Online]. Available: [http://link.springer.com/10.1007/978-3-642-31680-7\\_1](http://link.springer.com/10.1007/978-3-642-31680-7_1)
- [21] "PIPEDA Report of Findings #2018-003: Facebook agrees to stop using non-users' personal information in users' address books - Office of the Privacy Commissioner of Canada." [Online]. Available: <https://priv.gc.ca/en/opc-actions-and-decisions/investigations/investigations-into-businesses/2018/pipeda-2018-003/>
- [22] "News release: Privacy Commissioner launches Facebook investigation - Office of the Privacy Commissioner of Canada." [Online]. Available: [https://www.priv.gc.ca/en/opc-news/news-and-announcements/2018/nr-c\\_180320/](https://www.priv.gc.ca/en/opc-news/news-and-announcements/2018/nr-c_180320/)