OPEN ACCESS

Saudi Journal of Engineering and Technology

Abbreviated Key Title: Saudi J Eng Technol
ISSN 2415-6272 (Print) |ISSN 2415-6264 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com/sjeat

Original Research Article

Does Digital Privacy Really Exist? When the Consumer Is the Product

Bruno Silveira Cruz, MSc[1], Dr. Murillo de Oliveira Dias[2]*

[1]Fundação Getulio Vargas, Brazil
[2]Coordinator and Professor - Fundação Getulio Vargas, Brazil

DOI: 10.36348/sjet.2020.v05i02.00X | **Received:** 12.02.2020 | **Accepted:** 19.02.2020 | **Published:** X

***Corresponding author:** Murillo De Oliveira Dias

Abstract

In 2015, the scandal on Facebook and Cambridge Analytica Ltd, a British political consulting firm - subsidiary of the SCL Group, shook the international public opinion on digital privacy. The subject has attracted scholar attention, after 87 million mostly Facebook users worldwide, had their personal information under suspicion of data misappropriation, for political influence. In spite of the investigations conducted, a puzzling question remains: does digital privacy really exist? This article investigated the event and the role of the companies involved. Key findings point out that the sharing of personal identifiable information is a structured business model, with a vast ecosystem of providers and consumers. The article threw more light on digital privacy, and ultimately brought a full set of recommendations on data protection.

Keywords: Digital privacy, personal identifiable information, Facebook, Cambridge Analytica.

INTRODUCTION

The present case study investigated digital privacy, with Facebook and Cambridge Analytica Ltd data leakage case as the unit of analysis [1].

According to Facebook investors relations report, near 2.5 billion active users are detected on monthly basis [2]. In 2012, Facebook processed 2.5 billion pieces of content, 500 terabytes of data, 2.7 billion of “likes” and 300 million photos per day [3]. According to Ann Winbald, Facebook shareholder, “data is the new oil” [4].

Facebook also knows how to monetize on all its data. Until 2015, Facebook let its partners and developers that use Facebook APIs (Application Programming Interfaces) to access its users’ data and deliver focused ads to specific audiences, for instance [2]. The API was designed to collect users’ profiles, locations, likes and dislikes in Facebook, as well as the connections associated to contacts, in ever-increasing proportion. The objective is to portrait digital consumers’ preferences, that can be used, for instance, to offer commercial products focused on personal preferences. It is also possible to map political preferences, and therefore, to use data for political purposes too. This case investigated the case on Cambridge Analytica, a British political consulting firm - subsidiary of the SCL Group, shook the international public opinion on digital privacy, in 2015, with repercussions to date.

Data privacy represents, for digital social interactions, the same as secret vote, for democracies. Aleksandr Kogan, senior lecturer at Cambridge University, created a Facebook app called “this is your digital life”.

DIGITAL PRIVACY

Aleksandr then sent all collected data to Cambridge Analytica, without the final user’s consent [5].

METHODOLOGY

The present research is qualitative, interpretive, inductive reasoning, multiple-method approach, combining extensive archival research with descriptive case study, which unit is the case on data privacy Facebook-Cambridge Analytica, as the unit of analysis [1].

BACKGROUND

Launched on February 4th, 2004, by Mark Zuckerberg, Eduardo Severin, Andrew McCollum, Dustin Moskovitz and Chris Hughes, both Harvard under graduation students, “The Facebook”[1] took the world by storm.

[1] Later renamed just “Facebook”

One year before Facebook creation, Zuckerberg wrote the code for an experiment designed to attract attention and popularity, named *Facemash*. In the website, students from universities like Harvard could compare two photos from students, like the "hot or not" rating website. Maybe a predecessor for Tinder relationship website, with similar purposes.

Within the first hours of operations, *Facemash* attracted more than 450 visitors and 22,000 photo-views, breaking the data privacy of the Harvard database students. *Facemash* was later shutdown by Harvard administration, and Zuckerberg was charged with breach of security, violating copyrights and individual privacy. Zuckerberg recreated the initial idea of *Facemash* ahead of an art history final exam. Images were uploaded to a website where users could leave comments and share content with other users. A Face book is a student directory with photos and basic information. By 2003, Harvard University only had paper sheets distributed with students' directories and a few private online directories. Zuckerberg then created The Facebook. Initially in Harvard, very quickly expanded to Stanford, Columbia, Yale, and then to all Ivy league universities. By December 2005, Facebook had 6 million users.

On the other hand, Cambridge Analytica is a UK-based data analytics company. In 2015, Cambridge Analytica was incorporated by the SCL group, founded by Nigel Oakes. Cambridge Analytica aim is to predict and influence voter's behavior, though custom propaganda, specially crafted to impact a certain group or cluster of people that share similar characteristics. They do that by collecting data from multiple sources, analyzing and processing data with proprietary algorithms and statistic models in order to build specific profiles.

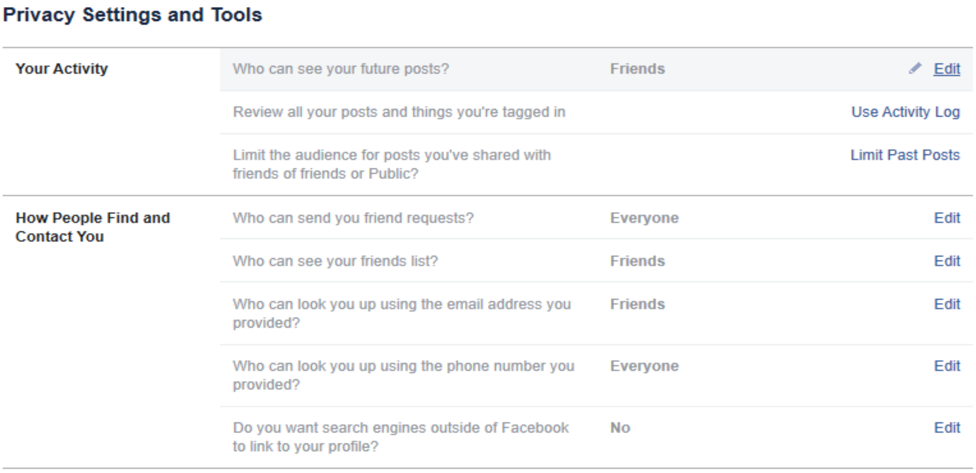
Although Aleksandr Kogan chose to pass data collected on Facebook to a third party (in case, Cambridge Analytica) without user's consent, it would be easy for Cambridge Analytica, or any other company or person for that matter, to do the same.

Facebook is, in fact, a marketing platform. Facebook's 2018 annual report for investors declared 55.8 billion revenue, in which 55 billion came from advertising campaigns [6]. It states explicitly: "We generate substantially all of our revenue from advertising. The loss of marketers, or reduction in spending by marketers could seriously harm our business" [2].

The same is valid for user engagement. If Facebook starts lagging on user engagement, their audience will start to shrink, pushing the price and volume of ads down, delivering a negative impact to the business. That's why Facebook is always launching new features and updating its platform. To keep users engaged, and to provide a qualified audience to marketers [2].

Therefore, by collecting users' data, Facebook creates and maintain qualified audiences. Everything a subscriber write, see, click or do, inside Facebook platform (which is comprised of Facebook, Instagram, Messenger, What's App, and a few others associated to Facebook). Data is then retrieved, organized, stored, combined, and processed in a myriad of different ways, to create customized profiles. [2] These profiles tell marketers what a given user likes and what dislikes, as well as their propensity to consume different products and services [2-4].

The process, however, is not transparent for the end user. By searching the settings on Facebook, one can access the *Privacy Settings* and *Tools* tabs. The users, thus, can choose whether friends are allowed to (i) check posts, (ii) send a friendship request, (iii) who friends list, among others, as depicted in the following Figure-1:

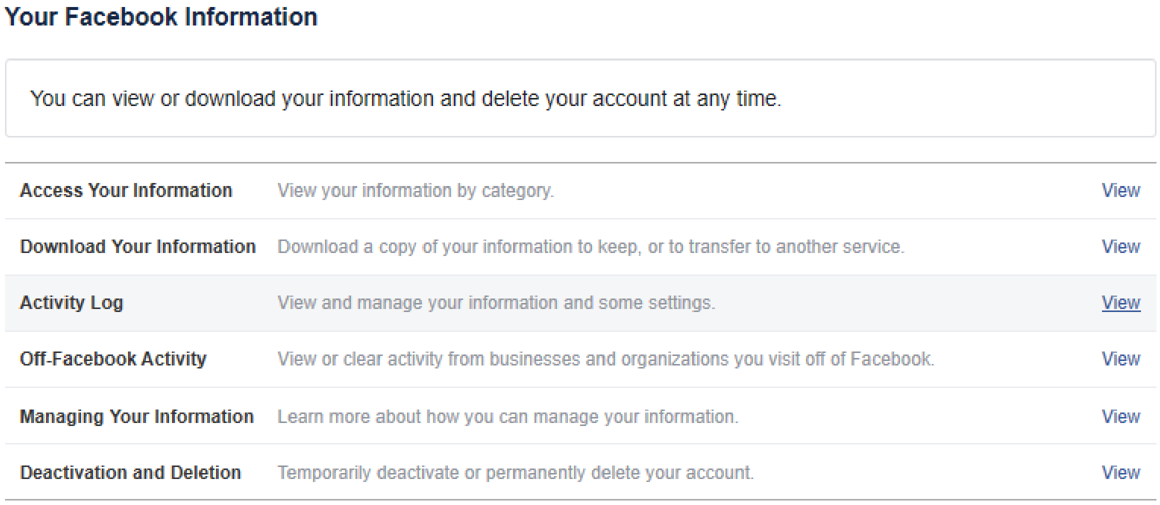


Privacy Settings and Tools

Your Activity	Who can see your future posts?	Friends	Edit
	Review all your posts and things you're tagged in		Use Activity Log
	Limit the audience for posts you've shared with friends of friends or Public?		Limit Past Posts
How People Find and Contact You	Who can send you friend requests?	Everyone	Edit
	Who can see your friends list?	Friends	Edit
	Who can look you up using the email address you provided?	Friends	Edit
	Who can look you up using the phone number you provided?	Everyone	Edit
	Do you want search engines outside of Facebook to link to your profile?	No	Edit

Fig-1: Privacy Settings and Tools
Source: Facebook, 2020

However, these options only scratch the surface of what is collected by Facebook. On the *Information* tab, depicted in Figure-2, a broader view of what kind of information is being collected, is illustrated:



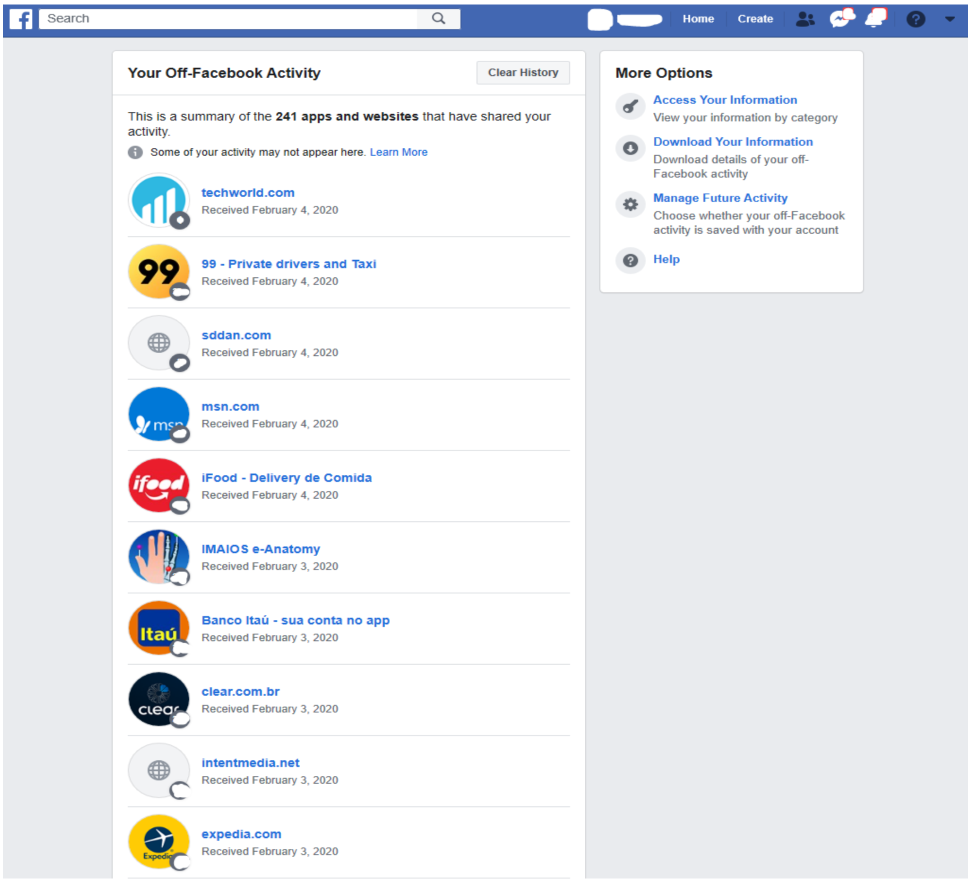
Your Facebook Information

You can view or download your information and delete your account at any time.

Access Your Information	View your information by category.	View
Download Your Information	Download a copy of your information to keep, or to transfer to another service.	View
Activity Log	View and manage your information and some settings.	View
Off-Facebook Activity	View or clear activity from businesses and organizations you visit off of Facebook.	View
Managing Your Information	Learn more about how you can manage your information.	View
Deactivation and Deletion	Temporarily deactivate or permanently delete your account.	View

Fig-2: You Facebook Information
Source: Facebook, 2020

When *Your Information* link is accessed, a wide range of categories appears. Every information one entered, uploaded or shared within Facebook is stored. From post, to videos, photos, places one has been, devices you use, everything one does is logged. Facebook also have a huge co-work network, who also collected and shared one data with Facebook, as depicted in Figure-3:



Your Off-Facebook Activity

Clear History

This is a summary of the **241 apps and websites** that have shared your activity.

Some of your activity may not appear here. Learn More

techworld.com
Received February 4, 2020

99 - Private drivers and Taxi
Received February 4, 2020

sddan.com
Received February 4, 2020

msn.com
Received February 4, 2020

iFood - Delivery de Comida
Received February 4, 2020

IMAIOS e-Anatomy
Received February 3, 2020

Banco Itaú - sua conta no app
Received February 3, 2020

clear.com.br
Received February 3, 2020

intentmedia.net
Received February 3, 2020

expedia.com
Received February 3, 2020

More Options

Access Your Information
View your information by category

Download Your Information
Download details of your off-Facebook activity

Manage Future Activity
Choose whether your off-Facebook activity is saved with your account

Help

Fig-3: Off Facebook Activity
Source: Facebook, 2020

Figure-3 shows the list of partners that had exchanged your information with Facebook. Observe the activities of a Facebook account. There are two things to be identified. First, the number of partners (companies) that had accessed your data is huge. In this case there were 241 partners that shared your activity with Facebook. In the list of partners there are many that will be instantly recognizable by the user, because the user uses a specific application on their devices, or because the user had visited a specific website. Second, in this list, there's also a few entries that do not seems to belong to the user. This happens because a technology called *cross-device tracking* [2, 3].

Cross-device Tracking is the technology used to track users on multiple devices [7]. Most sites and applications still use *cookies* to track users. A *cookie* file is created whenever you open an app or visit a website. It then stores your information and creates a unique identifier for your file, so you can be recognized the next time you visit the website or open the app. A cookie is stored on the device, so it doesn't matter which user is logged on the device, the cookie will track everything from its creation to its definitive deletion. As technology evolves, so do tracking techniques. The last advance in cross-device tracking uses human inaudible sounds to track users. Using audio beacons emitted by one device and recognized by the microphone of other device is possible to track what the user is seeing, and cross-track user activities based on device proximity. In short, Facebook established a cycle illustrated in the following Figure-4:

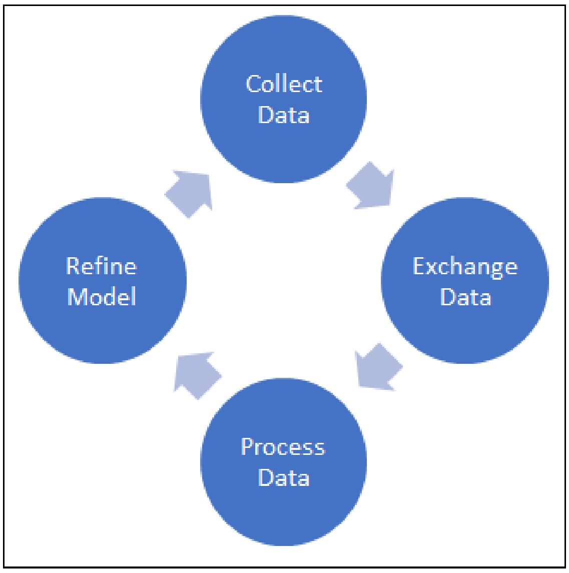


Fig-4: Facebook data cycle

CAMBRIDGE ANALYTICA

Cambridge Analytica was hired by Donald Trump's campaign for US presidency in 2016 to identify potential voters to be target of ads, and they also gave advice on how to impact those voters, where they were located and so on. Interestingly enough, Steve Bannon (Cambridge Analytica's board VP at that time) was also chosen to be Trump's chief strategist.

Later, Cambridge Analytica would also work to influence Brexit in the UK, at the same period. The company had won many contracts with politician eager to better impact their supporters and turn the tide on undecided voters. Barack Obama and Hillary Clinton's campaigns also used behavioral sciences to better reach specific audiences. Canadian company *AggregateIQ* was used by both candidates.

Both Cambridge Analytica and *AggregateIQ* are part of a large group of companies who use data to profile audiences. Other big names in this segment are Experian, FICO, Equifax and many others. Despite the similarities, those later companies work in a much dense battlefield of credit services, on top of other marketing initiatives. But the structure of data collection and processing, the creation and refinement of predictive models is the same, although each company use its own technologies and methodologies.

DISCUSSION

Public discussions are still going on about data privacy, big companies and small players operate between gray lines. Users are producing more and more data each single day, and it doesn't seem like this trend is going down anytime soon. Data can be used with all purposes, if they are not prevented to be use through specific legislation on the subject.

However, both society and governments are getting worried about the real power behind those companies. Is the population in general losing the power of choice, being directed to things they do not really want? Data privacy is so sensible that governments are acting: EU passed in 2018 the General Data Protection Regulation (GDPR). This law required that every company, whether establish in the EU or doing business with EU's companies or citizens, to apply certain standards when handling user data. It's the first real government action to protect users' data privacy [8].

Currently, is that toughest security and privacy law in the world. GDPR is an evolution of many years of privacy policies in the EU. The right of privacy was first established in 1950's European Convention of Human Rights. With the advance of the Internet, Europe was once again ahead passing the Data Protection Initiative in 1994. Finally, in 2016 GDPR passed on the European parliament and was put into effect. Companies were given two year to fully comply with the new law. GDPR is comprised of 11 chapters, 9 articles and 173 Recitals of Regulation. The law defines terms like personal data, data processing, data controller and many others, in order to leave no room for open interpretations. An in-depth analysis of the GDPR is out of the scope of this article, but one of the most important aspects of the law is the imposed limitations for companies and services providers to collect, store and process user data, without consent [8].

The law also enables anyone to withdraw previously consent at any time. And companies are obligated to honor the user's choice. Perhaps the most impactful thing about GDPR is the necessity to comply, even if the company is not in the EU. If you do business with any country within the EU, or any of its citizens, you must comply with the GDPR. However, there a few options for service providers to continue their business. Companies can collect, store and process user data without consent if, you need to comply with legal obligation, you're acting in public's interest or there's a legitimate interest, the last one being wide open to interpretation and discussion [8].

All the changes brought up by the GDPR, the Privacy Data Act (US) and the LGPD (Brazil), it is clear that the privacy options for service users will be changed dramatically. The Facebook options to control user privacy and data showed in this article is already a reflex of the limitations imposed by GDPR. [8] Now, users around the world need to understand, get to know and start to exercise their rights so their data is protected.

Finally, this study has implications in other business scenarios, such as: streaming video [9]; aircraft manufacturer industry [10]; e-business negotiation [11]; craft beer industry [12, 13], among others.

Future studies are encouraged to assess the impact of GDPR and similar laws, on the analytics market and whether both things can coexist, user privacy and analytics companies.

REFERENCES

1. Yin, R. (1988). Case Study Research: Design and Methods. Newbury Park, CA: Sage
2. Facebook Reports Fourth Quarter and Full Year 2019 Results. Retrieved from https://investor.fb.com/investor-news/press-release-details/2020/Facebook-Reports-Fourth-Quarter-and-Full-Year-2019-Results/default.aspx, on February 05, 2020.
3. Constine, J. (2012). How Big Is Facebook's Data? 2.5 billion Pieces of Content And 500+ Terabytes Ingested Every Day. TechCrunch. Retrieved from https://techcrunch.com/2012/08/22/how-big-is-facebooks-data-2-5-billion-pieces-of-content-and-500-terabytes-ingested-every-day/, on February 05, 2020.
4. Rotella, P. (2012). Is Data the New Oil? Forbes. Retrieved from https://www.forbes.com/sites/perryrotella/2012/04/02/is-data-the-new-oil/#259aecaa7db3, on February 05, 2020.
5. Sherr, I. (2018). Facebook, Cambridge Analytica and data mining: What you need to know. CNET. Retrieved from https://www.cnet.com/news/facebook-cambridge-analytica-data-mining-and-trump-what-you-need-to-know/, on February 05, 2020.
6. Facebook. (2018). Annual Report 2018. Retrieved from https://s21.q4cdn.com/399680738/files/doc_financials/annual_reports/2018-Annual-Report.pdf, on February 06, 2020.
7. Brookman, J. (2017). Cross-Device Tracking: Measurement and Disclosure. *Proceedings on Privacy Enhancing Technologies – Privacy Enhancing Technologies Symposium (PETS).* Retrieved from https://petsymposium.org/2017/papers/issue2/paper29-2017-2-source.pdf, on February 06, 2020.
8. Complete Guide to GDPR Compliance. (2020). Retrieved from https://gdpr.eu/, on February 08, 2020.
9. Dias, M., & Navarro, R. (2018). Is Netflix Dominating Brazil? *International Journal of Business and Management Review.* 6(1):19-32.
10. Cruz, B. S., Murillo de Oliveira, D. (2020). Crashed Boeing 737-Max: Fatalities or Malpractice? *Global Scientific Journals*. 8(1), 2615-2624.
11. Dias, M., & Duzert, Y. (2017). Teaching Materials: Role Play Simulation on E-Business Negotiation. *European Journal of Training and Development Studies*, 4(3), 1-15.
12. Murillo de Oliveira, D. (2020). Fatality, Malpractice, or Sabotage? Case on Craft Beer Poisoning in Minas Gerais, Brazil. *East African Scholars Multidisciplinary Bulletin*. 3(1), 26-31.
13. Murillo de Oliveira, D., & Teles, A. (2019). Facts and Perspectives on Craft Brewing Industry in Brazil. *International Journal of Management, Technology and Engineering*. 9(2), 1020-1028.