

FACEBOOK & CAMBRIDGE ANALYTICA'S DATA BREACH

By Vignesh Kumar and Liu Jason Tan

OUTLINE

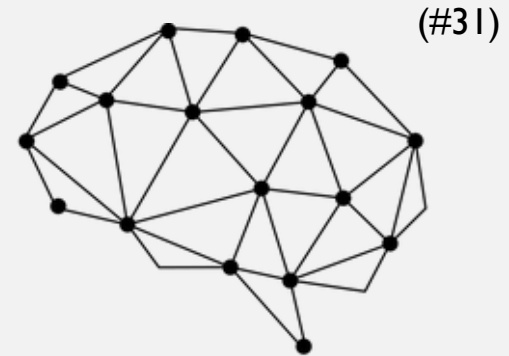
- Summary
- Cambridge Analytica
- Global Science Research
- Facebook
- FB's Open Graph API
- Open Graph capability
- “thisisyourdigitallife”
- FB's data policy changes
- FB's app authorization comparison
- Chris Wylie
- Initial aftermath
- Zuckerberg's response
- Zuckerberg's testimony before Congress
- Implications after Zuckerberg's testimony
- Cambridge University
- CA, SCL, Emerdata
- Implications of the event
- Significance
- Continuation of data privacy problems
- Similar issues

SUMMARY

- In 2018, news reports said approx. 87 million Facebook (FB) users had their data obtained by Cambridge Analytica (CA) without proper consent by an app that used the platform and was sold for political marketing for Trump's campaign
- A specific FB data access policy was taken advantage of by one of CA's affiliates
- Though Facebook was aware of the abuse in 2014 and changed data access policies as a result, they never enforced them
- A continuation of Facebook mishandling user data

CAMBRIDGE ANALYTICA

- A political data mining and analysis firm founded in 2013 at London, United Kingdom
 - Parent company - Strategic Communication Laboratories (now known as SCL Group)
- Co-founded by Robert Mercer and Steve Bannon
 - Robert Mercer
 - Republican mega donor
 - Steve Bannon
 - He heard about CA from a political strategist
 - Bannon convinced Mercer to invest in the firm and spoke to Trump about it during the campaign run
 - Became White House chief strategist for President Donald Trump
 - Left his position after 7 months of Trump's first year



Cambridge
Analytica

GLOBAL SCIENCE RESEARCH (GSR)

- A company co-founded by Aleksandr Kogan in 2014
 - A psychology researcher at the Cambridge University in the U.K.
- “Global Science Research optimizes marketing strategies with the power of big data and psychological sciences.” ⁽⁵⁾
- The deal with CA
 - CA paid GSR approx. \$800k in exchange for collected user data
 - They also paid an undisclosed amount of money to GSR for an unannounced product that was scrapped because of legal fees following the data breach scandal
 - Kogan says CA gave him no salary and his business went defunct



FACEBOOK

- A social media service that started in 2004
 - As of March 2019, 2.38 billion monthly active “users” worldwide
- Mark Zuckerberg is a co-founder and current CEO
- Social networks (and networks in general) can be represented as a graph data structure
 - Nodes/Vertices - users, Edges – relationships (i.e. friends)



FB'S OPEN GRAPH API

- Application programming interface (API) launched in 2010
 - Tools in software development to connect to a service
- App developers can access to an app user's data if that person authorizes it
 - “If accepted [by a user], “these apps would then have access to a user's name, gender, location, birthday, education, political preferences, relationship status, religious views, online chat status and more. With additional permissions, external sites could also gain access to a person's private messages.”
 - **At the time, developers could gain access to an app user's friends' data without their explicit consent**
 - FB let users know what data would be collected before using an app
- Zuckerberg wrote an opinion-editorial (op-ed) in *Washington Post* when this was released
 - He vowed to resolve to future user concerns about how their personal information was being managed by developers

CAPABILITY OF OPEN GRAPH

Permission Group	Permissions	Profile Items
Public profile (default)	public_profile	id, name, first name, last name, link, gender, locale, timezone, updated time, verified
App friends	user_friends	bio, birthday, education, first name, last name, gender, interested in, languages, location, political, relationship status, religion, quotes, website, work
Extended Profile Properties (xpP)	friends_about_me, friends_actions, friends_activities, friends_birthday, friends_checkins, friends_education_history, friends_events, friends_games_activity, friends_groups, friends_hometown, friends_interests, friends_likes, friends_location, friends_notes, friends_online_presence , friends_photo_video_tags, friends_photos, friends_questions, friends_relationship_details, friends_relationships, friends_religion_politics, friends_status, friends_subscriptions, friends_website, friends_work_history	about me, actions, activities, birthday checkins, history, events, games activity, groups, hometown, interests, likes, location, notes, online presence, photo video tags, photos, questions, relationship details, relationships, religion politics, status, subscriptions, website, work history
Extended Permissions (xP)	read_mailbox	inbox

FB APP AUTHORIZATION COMPARISON

(#38)

f Request for Permission

Yahoo! is requesting permission to do the following:

- Access my basic information**
Includes name, profile picture, gender, networks, user ID, list of friends, and any other information I've shared with everyone.
- Send me email**
Yahoo! may email me directly at [redacted].
Change
- Post to my Wall**
Yahoo! may post status messages, notes, photos, and videos to my Wall
- Access posts in my News Feed**
- Access my data any time**
Yahoo! may access my data when I'm not using the application
- Access Facebook Chat**
- Access my profile information**
Likes, Music, TV, Movies, Books, Quotes, About Me, Activities, Interests, Groups, Events, Birthday, Current City, Religious and Political Views, Education History, Work History and Facebook Status
- Access my friends' information**
Birthdays, Current Cities, Education History, Work History, Events and Facebook Statuses

By proceeding, you agree to the Yahoo! Terms of Service and Privacy Policy · Report App

Logged in as [redacted]

2010 (v1.0)

Allow **Don't Allow**

f Log in With Facebook

[Back](#) **Info You Provide**

Name and profile picture REQUIRED

Email address
test@example.com

Continue as [redacted]

Cancel

This doesn't let the app post to Facebook

[App Terms](#) · [Privacy Policy](#)

2019 (v2.9)

“THIS IS YOUR DIGITAL LIFE”

- A personality quiz hosted on Qualtrics
- Kogan (GSR) paid random people around \$3-4 to take his quiz
 - Amazon’s Mechanical Turk (MTurk) was used to distribute it (crowdsourcing)
 - He disclosed that it would be used for academic purposes
 - Approx. 270k FB users gave GSR access to their data
 - Paid approx. \$800k to them in total (the same amount that was given to him from CA)
- It requested access to certain parts of the app user’s data and that of their friends through the Open Graph API
 - Prior to 2014, there was no explicit consent needed for data access of app users’ friends
 - This explains the jump to 87 million affected users. App users’ friends were not aware that their data was collected and used for political purposes

(#37)



Continue with Facebook

FB'S DATA POLICY CHANGED

- In 2014, FB decided that third-party developers (i.e. Kogan) could no longer gain access to new data from an app user's friends
- In 2015, *The Guardian* posts an article that CA helped Ted Cruz's campaign by "psychographic profiling" ⁽⁴⁾
 - If you get information on what a person likes, what political party they support, where they live, and how old they are, you can create marketing that is curated to them
- FB responded to the article by banning "thisisyourdigitallife" from the platform and asked CA to remove data that was gained in violation of this policy
 - CA certified to FB that the data in question was indeed removed but not according to a whistleblower....

CHRIS WYLIE

- A whistleblower from CA
 - Former Director of Research for SCL Group and CA from 2013-2014
 - He left the company when CA's activities were "fracturing American society" ⁽³⁾
- Spoke to *The New York Times* and *The Guardian* in March 2018
 - Both publications released articles that brought the issue between CA and FB to the public
 - He freaked out when he saw his old boss Steve Bannon serve under Trump's administration
- CA's response to these articles
 - Denied using improperly obtained data during the 2016 elections as it was removed

(#35)



INITIAL AFTERMATH

- Users retaliated by deleting their FB accounts
- Federal Trade Commission (FTC) starts an investigation over FB
 - Wonders if the company violated a 2011 settlement agreement with FTC over data privacy
 - A \$3-5 billion settlement was recently made
- Various U.S. govt. officials inquire Zuckerberg to testify before Congress over misuse of data handling



(#36)

ZUCKERBERG'S RESPONSE

- Six days after the whistleblower articles came out, Zuckerberg apologizes
 - “A breach of trust” (Between FB and CA)
 - “We have a responsibility to protect your data, and if we can’t then we don’t deserve to serve you. I’ve been working to understand exactly what happened and how to make sure this doesn’t happen again.”
- New data policies were put in place (2018)
 - If a user no longer uses an app on FB for at least three months, developer access to new data from that user would be cut off
 - Any apps prior to the change in policy from 2014 are required to be audited by FB or be removed from the platform

ZUCKERBERG'S TESTIMONY BEFORE CONGRESS

- On April 10th, 2018, Mark Zuckerberg appeared before Congress to testify on behalf of Facebook. The following are quotes from Zuckerberg during his testimony
 - “We could have in theory banned [Cambridge Analytica] then (2015). We made a mistake by not doing so”
 - “What we allow is for advertisers to tell us who they want to reach, and then we do the placement ... That’s a very fundamental part of how our model works and something that is often misunderstood.”
 - “We’re investigating every single app that had access to a large amount of information in the past. And if we find that someone improperly used data, we’re going to ban them from Facebook and tell everyone affected.”
 - “It’s clear now that we didn’t do enough to prevent these tools from being used for harm. That goes for fake news, foreign interference in elections, and hate speech, as well as developers and data privacy.”

IMPLICATIONS AFTER ZUCKERBERG'S TESTIMONY

- The CEO said it would take three years to fix the problems Facebook is having with data and security and that the company has hired 15,000 of the 20,000 people to do so.
- To buy political or issue ads on Facebook, advertisers will have to verify their identity and location, and the company plans to roll out a tool that will let anyone see what ads a page is running. The company will also require people who manage large pages to verify their identity.
- Facebook has said that it plans to streamline its privacy settings and make it easier for people to understand what data Facebook is collecting.

CAMBRIDGE UNIVERSITY

- On April 11th, 2018, Zuckerberg said, “There's a whole program associated with the Cambridge University, where a number of researchers – not just Aleksandr Kogan, although to our current knowledge he's the only one who sold the data to Cambridge Analytica – there are a number of the researchers who are building similar apps.”
- Zuckerberg implied that the Kogan's project and collaboration with Cambridge Analytica is not isolated, and that there are similar situations at the University
- Cambridge University claims that there is no connection between the university and Cambridge Analytica

CA, SCL, AND EMERDATA

- On May 1st, 2018, CA files for bankruptcy and announces that it and SCL Group would close
 - Though they deny any responsibility following the data breach, legal fees from investigations and negative media coverage brought them down
- Emerdata
 - Formed in August of 2017 by SCL Group executives but was not incorporated into it
 - Part of Firecrest Technologies Ltd. which is a separate company
 - After the scandal broke, the company changed its board of directors and received new funding
 - Nigel Oaks, the founder of SCL, said they had plans for Emerdata to acquire CA & SCL (initially), however, many employees of CA left following the scandal, so it was not worth to incorporate it
 - It's questionable to whether CA's operations are continuing here

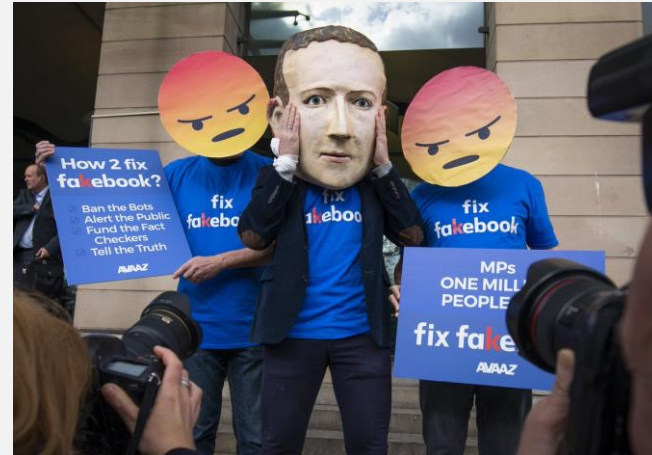
IMPLICATIONS OF THE EVENT

- Mark Zuckerberg said, “new regulations are needed in four areas: harmful content, protection of elections, privacy and data portability.”
- The European Union passed a new regulation called **General Data Protection Regulation (GDPR)** for data protection and privacy
 - There are eight rights for individuals. These include allowing people to have easier access to the data companies hold about them, a new fines regime and a clear responsibility for organisations to obtain the consent of people they collect information about.
- California State Legislature passed the **California Consumer Privacy Act (CCPA)** for privacy rights and consumer protection. It is to be in effect January 1st, 2020. Key highlights include:
 - Right to know all data collected by a business on you
 - Right to say no to the sale of your information
 - Right to sue companies who collected your data, where that data was stolen or disclosed pursuant to an unauthorized data breach
 - Right to delete the data you’ve posted

SIGNIFICANCE

- Alongside the CA-FB breach, there were these problems in social media at the time:
 - A successful Russian disinformation campaign that undermined the 2016 U.S. elections
 - Concerns for its role on mental health
- The problem with social media is trust
 - A user's relationship with that content is meaningful and has value
 - In the case of the breach, it was mishandled by FB through their 2010 data access policy.
 - CA took advantage of this and FB never regulated/enforced its policy changes.
 - FB failed to protect its users' privacy

(#39)



CONTINUATION OF FACEBOOK DATA PRIVACY PROBLEMS

- Reportedly, Zuckerberg leveraged user data with various competitors and partners as shown through leaked, internal documents (April 16, 2019)
 - They would limit or allow more access to user data with other companies depending on their relationship with them; it was used as a bargaining chip
 - The company owns WhatsApp and noticed a different messaging app owned by Yahoo called MessageMe (now shutdown) trending, so FB restricted what those developers could take from the Open Graph API
- FB's mishap with email contacts (April 18th, 2019)
 - 1.5 million email contacts were mistakenly uploaded to FB servers
 - It was part of a feature where these contacts were used to find friends on the platform
 - These contacts are now in the process of being deleted.

SIMILAR ISSUES

- CubeYou, a data analytic firm similar to Cambridge Analytica, used personality quizzes, similar to 'thisisyourdigitallife' to help marketers find customers (April 2018)
 - Facebook said they are going to suspend CubeYou from the platform to investigate and CubeYou is seeking reinstatement
- Twitter urged all of its users to immediately change their passwords after a bug exposed them in plain text (May 2018)
 - Twitter's investigation showed that there was no evidence that any breach or misuse of the unmasked passwords occurred; recommendation out of an "abundance of caution,"
- In late 2016, data from Uber users and drivers were leaked
 - Two hackers were able to get names, email addresses, and mobile phone numbers of 57 million users and the driver license number of 600,000 drivers.

REFERENCES

1. <https://www.cnn.com/2018/04/10/facebook-cambridge-analytica-a-timeline-of-the-data-hijacking-scandal.html>
2. <https://www.vox.com/policy-and-politics/2018/3/23/17151916/facebook-cambridge-analytica-trump-diagram>
3. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>
4. <https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data>
5. <https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie>
6. <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html>
7. <https://www.forbes.com/sites/kathleenchaykowski/2018/03/21/mark-zuckerberg-addresses-breach-of-trust-in-facebook-user-data-crisis/#377f79e23e36>
8. <https://www.marketwatch.com/story/here-are-the-changes-facebook-has-announced-ahead-of-zuckerberg-testimony-2018-04-09>
9. <https://www.theguardian.com/technology/2018/apr/11/mark-zuckerbergs-testimony-to-congress-the-key-moment>
10. https://en.wikipedia.org/wiki/Cambridge_Analytica
11. <https://techcrunch.com/2015/04/28/facebook-api-shut-down/>
12. <https://www.theguardian.com/news/2018/apr/24/aleksandr-kogan-cambridge-analytica-facebook-data-business-ventures>
13. <https://www.theguardian.com/uk-news/2018/may/02/cambridge-analytica-closing-down-after-facebook-row-reports-say>
14. <https://www.theguardian.com/uk-news/2018/may/05/cambridge-analytica-scl-group-new-companies-names>

REFERENCES (CONT.)

15. <https://www.theverge.com/2017/12/15/16781448/facebook-makes-you-feel-bad-study-research>
16. <https://money.cnn.com/2018/03/18/technology/business/facebook-cambridge-analytica/index.html>
17. <https://www.fastcompany.com/90231904/a-facebook-employee-tied-to-cambridge-analytica-quietly-left-facebook>
18. <https://www.cbsnews.com/news/aleksandr-kogan-the-link-between-cambridge-analytica-and-facebook-60-minutes/>
19. <https://www.crunchbase.com/organization/global-science-research#section-overview> (image)
20. <https://www.nbcnews.com/tech/social-media/mark-zuckerberg-leveraged-facebook-user-data-fight-rivals-help-friends-n994706>
21. <https://www.businessinsider.sg/facebook-uploaded-1-5-million-users-email-contacts-without-permission-2019-4/?r=US&IR=T>
22. <https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018>
23. <https://www.caprivacy.org/about>
24. <https://www.politico.com/story/2019/04/02/facebook-ftc-fine-1271618>
25. <https://thetab.com/uk/cambridge/2018/04/13/how-is-cambridge-university-linked-to-cambridge-analytica-and-the-facebook-data-scandal-110205>
26. <https://www.cNBC.com/2018/04/08/cubeyou-cambridge-like-app-collected-data-on-millions-from-facebook.html>
27. <https://www.theverge.com/2018/5/3/17316684/twitter-password-bug-security-flaw-exposed-change-now>
28. <https://www.nytimes.com/2018/09/26/technology/uber-data-breach.html>

REFERENCES (CONT.)

29. <https://www.ft.com/content/b1a49898-2c44-11e8-a34a-7e7563b0b0f4>
30. <https://eprint.iacr.org/2015/456.pdf> (image)
31. <https://www.fastcompany.com/40550423/how-facebook-blew-it>
32. <https://investor.fb.com/home/default.aspx>
33. By Source (WP:NFCC#4), Fair use, <https://en.wikipedia.org/w/index.php?curid=57318211> (image)
34. https://www.iconfinder.com/icons/313103/facebook_fb_icon (image)
35. <https://www.timescolonist.com/news/local/how-a-victoria-kid-ended-up-at-heart-of-facebook-data-mining-story-1.23206550>
(image)
36. <https://techcrunch.com/2018/03/19/deletefacebook/> (image)
37. <https://slickdeals.net> (image)
38. <https://www.blogsolute.com/access-facebook-profile-inside-yahoo-mail/11081/> (image)
39. <https://twitter.com/Avaaz/status/98946676722296579> (image)

QUESTIONS?