Episode #260
The Cambridge Analytica Scandal
6th May, 2022

[00:00:00] Hello, hello hello, and welcome to English Learning for Curious Minds, by Leonardo English.

[00:00:12] The show where you can listen to fascinating stories, and learn weird and wonderful things about the world at the same time as improving your English.

[00:00:22] I'm Alastair Budge, and today we are going to be talking about The Cambridge Analytica **Scandal**[1].

[00:00:29] It's a story that brings together the power of Facebook, journalism, political **prejudice**[2], right vs. left, democracy, human psychology, free will, **allegations**[3] of

---

[1] an event that caused a public feeling of shock and anger

[2] a wrong and unfair opinion or feeling not based on reason or actual experience

[3] suggestions that someone had done something illegal or wrong without having proof

Russian spies working to influence foreign elections and a **plot**[4] to put Donald Trump in the White House.

[00:00:49] The only problem is.

[00:00:50] Is any of it even true? Right, let's get started and talk about Cambridge Analytica.

[00:01:00] If you had bought the Guardian Newspaper on March 18th of 2018, on the front cover you would have seen a picture of a 28-year-old man with short, pink hair.

[00:01:14] Next to him, the **all-powerful**[5] headline:

[00:01:18] "Revealed: 50 million Facebook files taken in record data **breach**[6]."

[00:01:25] A **breach**, by the way, means a break, a hole that has been made in order to get in and access something.

[00:01:34] The story went on to reveal that a company called Cambridge Analytica had used the personal Facebook data of 50 million people to target US citizens, build a–and I'm quoting directly here–"psychological warfare tool" which targeted people based on

---

[4] secret plan

[5] extremely powerful

[6] a break, a hole that has been made in order to get in and access something

their stolen data, and ultimately **swing**[7] the election, to win the 2016 US presidential

election for Donald Trump.

[00:02:04] It was a powerful story, and it was all over the news.

[00:02:09] You may well remember the story yourself.

[00:02:12] It created big problems for Facebook, and the company's share price. It lost

$35 billion in value within a day, and in the **subsequent**[8] months it lost over $100 billion

in stock market value.

[00:02:29] The company behind the supposed data **breach**, Cambridge Analytica, **went

out of business**[9], and its former CEO was forced to **testify**[10] in front of the British

parliament.

[00:02:43] The journalist behind the story, Carole Cadwalladr, was invited to give a TED

talk, and won the **prestigious**[11] George Orwell prize for journalism for breaking this

story.

---

[7] change from one opinion to another

[8] following

[9] closed down, stopped operating

[10] give information or evidence in a law court

[11] showing respect and admiration

[00:02:56] So, what we are going to try to answer in this episode is what actually happened?

[00:03:03] Is democracy really at threat from bad actors using personal data to influence our voting behaviour?

[00:03:11] Or is the entire story completely **overblown**[12]?

[00:03:15] So, let's start with what happened.

[00:03:19] The main character in the story, without whom none of this would have been possible, was a man called Aleksandr Kogan.

[00:03:27] He was born in the USSR, the former Soviet Union, but moved to the United States when he was 7 years old.

[00:03:35] Kogan **excelled**[13] at school, and showed a **remarkable**[14] talent for mathematics and physics.

[00:03:43] At university he became increasingly interested in psychology, in understanding why humans feel certain emotions, and what makes us behave in the way we do.

---

[12] made to seem more impressive or important than it is

[13] was very good

[14] worthy of attention, amazing

[00:03:56] He had been working at Cambridge University since 2012, and saw from early on the potential power that a then **newish**[15] technology company, Facebook, had. If you used Facebook back in 2012, I imagine it played a bigger, or at least more obvious, **role**[16] in your life than it does now.

[00:04:20] You used that "like" button to like anything from a status update from a friend to a "raising money for dog shelters" campaign.

[00:04:29] You probably didn't think too much about what happened after you clicked that "Like" button. After all, Facebook was a fun way to keep in touch with friends and family, share photos and generally see what was going on in the world.

[00:04:46] To Kogan, however, it seemed like the most incredible dataset on humanity.

[00:04:53] People like you and me were spending all this time on Facebook, liking stuff, **interacting**[17] with it, giving Facebook data on what we like and what we don't like, what we **engage**[18] with and what we aren't interested in.

[00:05:08] What's more, Facebook could see who all of our friends were, so it had this amazing understanding of connections between all of its users.

---

[15] somewhat new, recent

[16] part

[17] reacting to

[18] occupy ourselves or are involved in

[00:05:18] In 2014, Facebook passed 1.3 billion users, each of whom was spending an average of 40 minutes a day on the network.

[00:05:30] So not only did it have a vast amount of people using it, but they were using it a lot.

[00:05:37] To someone like Kogan, this was fascinating.

[00:05:41] What's more, at the time, Facebook had recently started to allow **third-parties**[19] to develop applications on top of Facebook that would give them certain permissions to view the data on Facebook users.

[00:05:56] If you can think back to this time, if you were a Facebook user that is, you might remember a **load**[20] of quizzes where you **granted**[21] the quiz access to your Facebook account, answered a few questions, and it would tell you things like what Harry Potter character you were like or what member of the Beatles you would have been.

[00:06:16] Silly stuff, **harmless**[22] fun. Or so most people thought.

---

[19] groups outside facebook

[20] lot

[21] gave

[22] not able to cause damage

[00:06:21] Indeed, back in 2014 these **third-party**[23] applications were able to access some data not just about you but also about your friends.

[00:06:32] Someone might give access to a quiz to find out what type of pizza they are, agreeing to share their data with the app, but not realise that they were actually allowing the application to access information about their friends as well.

[00:06:49] Kogan's question was, was any of this Facebook data actually any good at telling you anything useful about that person's personality?

[00:06:59] If someone on Facebook liked "dogs, Mars bars and reading", did this actually tell you anything about them as an individual?

[00:07:10] But Kogan was no **thief**[24], and he couldn't just take this data from Facebook. Facebook users needed to give him their permission to use it.

[00:07:20] And it's here that we meet Cambridge Analytica.

[00:07:24] Cambridge Analytica, which has no links whatsoever to Cambridge University, by the way, was a political consulting company founded in 2013.

[00:07:35] Its supposed speciality was advising political campaigns, advising politicians on how to use social media and data to put the right message in front of the right people at the right time.

---

[23] belonging to groups outside facebook

[24] a person who steals, who takes something without permission

[00:07:50] Cambridge Analytica paid Kogan to create an app that would collect data on Facebook users.

[00:07:57] Cambridge Analytica would be able to use the data for use with its political clients, and Kogan could continue his academic research with the data he collected.

[00:08:08] A win-win situation, one where both parties benefited.

[00:08:13] So, Kogan, with Facebook's permission, built a simple application called "This Is Your Digital Life".

[00:08:21] It was a quiz, essentially.

[00:08:23] Users were offered three or four dollars to participate in the survey, and in exchange they gave the app permissions to access their Facebook data.

[00:08:35] What they probably didn't realise, though, was that they were actually giving the app permission to see some data on all of their friends.

[00:08:43] So if I **authorised**[25] the app, if I took the quiz, the app would get data on all of my friends. If I had 500 friends, for example, the app got the data on those 500 friends.

[00:08:58] The quiz was taken by around 300,000 people, each one with hundreds of friends, meaning that Kogan now had Facebook data on almost 90 million people. The initial Guardian report said 50 million, but Facebook was later forced to admit that it had actually been almost double this number.

---

[25] gave permission to

[00:09:19] Getting this data wasn't cheap, as the app had to **incentivise**[26] people to take the personality test. Cambridge Analytica provided the financing for it, about a million dollars, believing that the data that would come out at the end of it would be significantly more valuable.

[00:09:38] Now, let's just pause to address a couple of points first, before we move on with the story.

[00:09:44] Kogan had Facebook's permission to create the quiz application. There's no **dispute**[27] about that.

[00:09:51] What is disputed is that Kogan says he was given permission to sell this data on to **third-parties**. Facebook says he wasn't.

[00:10:01] Where it does become a little **blurry**[28], a little less clear, is about how the data on Facebook users was accessed, and the use of the term "**breach**".

[00:10:12] As a reminder, a **breach** is when illegal access is provided to something.

[00:10:18] At the time, Facebook's **default**[29] settings meant that if you gave an app permission to your data, by default it could get data on all of your friends.

---

[26] give them reasons

[27] disagreement, argument

[28] not clear

[29] preset or pre-selected

[00:10:29] The Facebook users likely didn't understand exactly what they were giving their permission for, but Kogan's app was one of hundreds, thousands even, that was collecting very similar amounts of data.

[00:10:43] So Kogan wasn't doing anything illegal, and indeed he was doing what practically every other Facebook app was doing.

[00:10:52] Ok, so with those **clarifications**[30] out of the way, let's return to our story.

[00:10:58] You might be thinking, what is Kogan actually getting out of all of this?

[00:11:03] Was he really a Russian agent, was he a **closet**[31] Trump supporter, did he want to **wreak havoc**[32] in foreign elections and provide data that could be used to **undermine**[33] democracy?

[00:11:15] That would make for a good story, but unfortunately it doesn't seem to be true.

[00:11:21] He says that he was **unaware**[34] that the data would be used for political targeting, and his interest in doing all of this was primarily from a research perspective.

---

[30] actions of making the situation less confusing

[31] secret

[32] cause great damage

[33] weaken, harm

[34] having no knowledge

[00:11:33] He wanted to know whether all of this data he was collecting on people - what they liked, where they lived, what they did, whether this data actually helped **predict**[35] anything about their personalities.

[00:11:46] Remember, he is an amazingly talented scientist with an interest in human psychology. He wanted to know how human behaviour and personality can be predicted.

[00:11:58] And as for Cambridge Analytica, if Kogan's work proved to be able to predict human behaviour and personality, it would be amazingly valuable.

[00:12:08] Imagine that just by knowing that if someone liked Mars bars, dogs and reading they would be particularly **receptive**[36] to a certain type of political message, and if someone liked singing, learning languages and had a birthday in December then they would be **receptive** to another type of political message.

[00:12:29] **Armed**[37] with this information on people, and with Facebook, the ability to **target**[38] over a billion people, and the vast majority of the voters in the United States, this would be an incredibly powerful weapon.

[00:12:44] But how could you actually figure out whether the data was useful or not?

---

[35] know or tell in advance

[36] willing to consider or accept

[37] equipped, provided

[38] have as objects of attention

[00:12:49] Well, that's where the personality quiz came in.

[00:12:54] After someone gave access to Kogan's application, they needed to **fill out**[39] the personality quiz.

[00:13:01] Sure, you can say that people aren't very good at **assessing**[40] their own personality, but as there was no reason for someone to fill in a question incorrectly, and the test was completely randomised, this data should have been accurate.

[00:13:17] After the Facebook user had completed the quiz, Kogan's **algorithm**[41] looked at their answers and classified them by five different **personality traits**[42] - such as if they were an **extrovert**[43] or **introvert**[44], how open they were to new ideas, and so on.

[00:13:34] It was then time to see whether all the data he had collected from Facebook was actually any good at predicting these people's personality.

[00:13:44] It was **crunch time**[45].

---

[39] complete the form with the information asked

[40] judging, evaluating

[41] a set of rules that makes it possible for a computer to solve a problem or perform a task

[42] personality characteristics

[43] an outgoing, sociable person

[44] shy, not sociable

[45] a period of pressure, a critical situation

[00:13:46] Unfortunately, according to Kogan, the **algorithms**[46] managed to correctly predict someone's personality 1% of the time. That is 1%, 1 one out of 100 times, meaning that 99% of the time it got it wrong.

[00:14:04] The Facebook data simply didn't seem to be any good at all at predicting someone's personality.

[00:14:11] Now, you might have thought that this would be the end of the story, with Cambridge Analytica **cutting its losses**[47] as the data was **junk**[48].

[00:14:20] But really it was just the start.

[00:14:23] Cambridge Analytica seemed **utterly**[49] **uninterested**[50] in how accurate the data actually was.

[00:14:30] Instead, it continued to tell its clients that it had an incredibly valuable **dataset**[51] that could be used for political **targeting**[52].

---

[46] sets of rules that make it possible for a computer to solve a problem or perform a task

[47] stopping the activity in order to not to lose more money

[48] useless

[49] completely

[50] not interested

[51] collection of information that can be used by a computer

[52] the act of selecting something as an object of attention

[00:14:39] To be precise, it said it had up to 5,000 data points on over 220 million Americans.

[00:14:48] The company also had connections in **high places**[53]. Its investors included the billionaire Republican **donor**[54], Robert Mercer, and the man who would later go on to become Donald Trump's chief **strategist**[55], Steve Bannon.

[00:15:03] These powerful connections would give it a **head start**[56] in the US political scene, and would later result in it working on the 2016 US presidential campaign.

[00:15:15] While it would be most famous for working with Donald Trump, its first client was actually his competitor, the Republican **establishment**[57] candidate, Ted Cruz.

[00:15:26] Cambridge Analytica had managed to sell Cruz on the power of its data and analytics **capabilities**[58], but as you will know, Cruz was heavily **beaten**[59] by Trump in the nomination to be the Republican candidate for president.

---

[53] positions of power or authority

[54] financial supporter

[55] someone who planned his political matters

[56] advantage

[57] supported by the established or more powerful political group or regime

[58] abilities or potential

[59] defeated

[00:15:42] Sure, you might think, data can't **compensate**[60] for an **uninspiring**[61] candidate, and there was only a limited amount that Cambridge Analytica could have done for Cruz.

[00:15:53] In any case, it didn't work.

[00:15:56] But that didn't stop the company from being employed by the eventually **victorious**[62] **anti-establishment**[63] Trump campaign in 2016.

[00:16:06] It later **emerged**[64] that Cambridge Analytica had told potential clients that it had also consulted on the Brexit campaign, the campaign to leave the EU in June of 2016, which happened six months before Trump's victory in November.

[00:16:23] And the CEO of Cambridge Analytica, a tall man named Alexander Nix with an almost Bond-villain style, would present at conferences and give interviews where he talked about the power of the company's data and analytics **capabilities**, **boasting**[65]

---

[60] fix or correct the situation

[61] not producing excitement or interest

[62] having won

[63] not supported by the established or most powerful political group or regime

[64] became known

[65] speaking too proudly

about how many academics and PhDs it employed, and how it was able to **harness**[66]

the power of big data to change public opinion about anything.

[00:16:52] Then, in March of 2018, the **scandal broke**[67].

[00:16:57] It was front-page news in The Guardian and the New York Times, and was all

over cable news around the world the very same day.

[00:17:06] For everyone who wasn't involved in the world of political campaigning, so

for almost everyone, it was the first time they had heard the name Cambridge

Analytica.

[00:17:16] And it was frightening.

[00:17:18] Experts were invited to come on TV and talk about this **threat**[68] to

democracy.

[00:17:24] Could free will continue to exist if companies were able to predict and change

our voting patterns without our knowledge?

[00:17:33] For many, it was the first time that they had thought about the power and

possibilities of the data that Facebook had on them.

---

[66] control

[67] suddenly began

[68] danger

[00:17:42] The hashtag #DeleteFacebook was trending on Twitter, and Mark Zuckerberg was **dragged**[69] to **testify** before the US Congress while his company was losing billions of dollars of value.

[00:17:55] It made for a wonderful story.

[00:17:58] To people who had questioned how a country could possibly have voted for Brexit or Donald Trump, it provided an explanation.

[00:18:07] The people had been **tricked**[70], psychologically profiled and targeted with messages to influence their opinion.

[00:18:15] There was even the fact that Alexander Kogan was born in Russia, which to some was a sign that he was a Russian spy. Journalists called him up and asked him **point-blank**[71] whether he was a Russian agent.

[00:18:29] Suddenly it was clear, and everything made sense.

[00:18:33] But the reality is that the **scandal** was a bit of **a storm in a teacup**[72], it was a **scandal** over something that didn't really happen.

---

[69] caused to be involved in an unpleasant situation

[70] fooled, deceived

[71] directly and openly

[72] a situation in which people were very angry or upset about something that didn't really happen

[00:18:43] Kogan, the man who collected the data and ran the statistical models, has said numerous times that the data was practically **useless**[73] on a personal level, even when it was first collected.

[00:18:57] What's more, it was collected in 2014 and would have been pretty **out of date** [74] in 2016 anyway. Data like this typically has a **shelf-life**[75] of a maximum of 12 months.

[00:19:11] Plus, the fact that the Cruz campaign and the Trump campaign both got rid of Cambridge Analytica suggests that the people who paid to use the data, and supposedly used it to target voters with **hyper-targeted**[76] advertising and swing an election, didn't find it useful at all.

[00:19:31] And as far as the question of whether the data had any impact on the Brexit vote, there was an **extensive**[77] enquiry by the UK Information Commissioner's Office, essentially an arm of government, which found absolutely no evidence that Facebook data held by Cambridge Analytica had any influence on the Brexit vote.

[00:19:53] So, let's **recap**[78] quickly, as there are two slightly different questions:

---

[73] of no use

[74] no longer relevant or accurate

[75] the length of time for which it remains usable

[76] aims (or directs its attention towards) advertising messages to a very specific group of people

[77] large in scale, in-depth and detailed

[78] repeat the main points

[00:19:58] Firstly, was there an illegal **breach** of Facebook data?

[00:20:03] And secondly, did this data actually influence anything?

[00:20:08] To the first question, no, it wasn't illegal, but it was probably **misleading**[79]. The users who **authorised** this application probably didn't know that they were also giving their friends' data, or that it might be used for commercial purposes.

[00:20:25] As a result, Facebook has **clarified**[80] its policies on this, and **third-party** applications don't get access to your friends' data, and they get access to much less data than they used to get.

[00:20:39] If you **authorise**[81] a **third-party** app now on Facebook, it's a lot clearer, and these apps don't **tend**[82] to get access to very much.

[00:20:49] And to our second question, did the data actually influence anything, although it makes for a much **juicier**[83] and more interesting story if you think it does, the reality certainly appears to be "no".

---

[79] giving the wrong idea or impression

[80] made them clear

[81] give permission to

[82] usually do it

[83] causing more curiosity and interest

[00:21:04] Now, as a final point, just in case this is mistaken for some great defence of Facebook or Cambridge Analytica, it most definitely isn't.

[00:21:14] The idea that you have accidentally given access to your personality, what you like and dislike, who you are, the idea that some evil **third-party** knows it is scary. It's not nice, and it's definitely not nice to think that you might have let this third-party have access to your **nearest and dearest**[84].

[00:21:35] But the reality is that this data, at least in the case of the Cambridge Analytica **scandal**, is not nearly as useful as you might think it to be.

[00:21:45] Indeed, there is one theory that the Cambridge Analytica **scandal** was actually excellent news for Facebook, because it made the world think that the data Facebook had on its users was significantly more valuable than Kogan's analysis suggested it actually was.

[00:22:03] Now, to go back to our example of someone who likes "dogs, Mars bars and reading".

[00:22:10] Well, I like "dogs, Mars bars and reading".

[00:22:13] Maybe you like "dogs, Mars bars and reading" too.

---

[84] close friends and relatives

[00:22:17] Does this mean that we're similar, and most importantly, do these three things mean that we are likely to have the same political beliefs, care about the same issues, and are likely to respond in the same way to a political advert?

[00:22:32] It's a possibility, but from what Kogan's research suggests, it certainly seems pretty unlikely.

[00:22:40] This is, of course, not to say that the information Facebook holds on you or me is **worthless**[85] and **unusable**[86].

[00:22:47] Facebook knows a lot more about us than whether we like dogs, Mars bars and reading, and hundreds of millions of dollars are spent every single day advertising to users on Facebook because of what Facebook knows about them.

[00:23:02] But can this data be easily stolen by a **third party**[87], or even a foreign agent, and used to **manipulate**[88] democracy?

[00:23:12] Well, with the example of Cambridge Analytica, it certainly seems that even when we think it has, it really hasn't.

[00:23:22] OK then, that is it for today's episode on the Cambridge Analytica **Scandal**.

---

[85] having value or worth

[86] cannot be used

[87] a group or organisation less directly involved

[88] control in an unfair or dishonest way

[00:23:28] I hope it's been an interesting one, that you've learnt something new, and if you remember this **scandal** from a few years ago, well perhaps it has put a different perspective on it.

[00:23:39] As always, I would love to know what you thought of this episode.

[00:23:42] Do you remember this news coming out? How did you feel when it did?

[00:23:47] If you are a Facebook user, or I should probably say Meta now, shouldn't I, anyway, how did it affect your behaviour, if at all?

[00:23:56] I would love to know, so let's get this discussion started.

[00:24:00] You can head right into our community forum, which is at community.leonardoenglish.com and get chatting away to other curious minds.

[00:24:09] You've been listening to English Learning for Curious Minds, by Leonardo English.

[00:24:14] I'm Alastair Budge, you stay safe, and I'll catch you in the next episode.

[END OF EPISODE]

Key vocabulary

Word	Definition
Scandal	an event that caused a public feeling of shock and anger
Prejudice	a wrong and unfair opinion or feeling not based on reason or actual experience
Allegations	suggestions that someone had done something illegal or wrong without having proof
Plot	secret plan
All-powerful	extremely powerful
Breach	a break, a hole that has been made in order to get in and access something
Swing	change from one opinion to another
Subsequent	following
Went out of business	closed down, stopped operating
Testify	give information or evidence in a law court
Prestigious	showing respect and admiration

Overblown   made to seem more impressive or important than it is

Excelled   was very good

Remarkable   worthy of attention, amazing

Newish   somewhat new, recent

Role   part

Interacting   reacting to

Engage   occupy ourselves or are involved in

Third-parties   groups outside facebook

Load   lot

Granted   gave

Harmless   not able to cause damage

Third-party   belonging to groups outside facebook

Thief   a person who steals, who takes something without permission

Authorised   gave permission to

Incentivise   give them reasons

Dispute   disagreement, argument

Blurry   not clear

Default   preset or pre-selected

Clarifications   actions of making the situation less confusing

Closet   secret

Wreak havoc   cause great damage

Undermine   weaken, harm

Unaware   having no knowledge

Predict   know or tell in advance

Receptive   willing to consider or accept

Armed   equipped, provided

Target   have as objects of attention

Fill out   complete the form with the information asked

Assessing   judging, evaluating

Algorithm   a set of rules that makes it possible for a computer to solve a problem

                       or perform a task

Personality traits   personality characteristics

Extrovert   an outgoing, sociable person

Introvert   shy, not sociable

Crunch time   a period of pressure, a critical situation

Algorithms   sets of rules that make it possible for a computer to solve a problem or

perform a task

Cutting its losses   stopping the activity in order to not to lose more money

Junk   useless

Utterly   completely

Uninterested   not interested

Dataset   collection of information that can be used by a computer

Targeting   the act of selecting something as an object of attention

High places   positions of power or authority

Donor   financial supporter

Strategist   someone who planned his political matters

Head start   advantage

Establishment	supported by the established or more powerful political group or regime
Capabilities	abilities or potential
Beaten	defeated
Compensate	fix or correct the situation
Uninspiring	not producing excitement or interest
Victorious	having won
Anti-establishment	not supported by the established or most powerful political group or regime
Emerged	became known
Boasting	speaking too proudly
Harness	control
Broke	suddenly began
Threat	danger
Dragged	caused to be involved in an unpleasant situation
Tricked	fooled, deceived

Point-blank   directly and openly

A storm in a teacup   a situation in which people were very angry or upset about something that didn't really happen

Useless   of no use

Out of date   no longer relevant or accurate

Shelf-life   the length of time for which it remains usable

Hyper-targeted   aims (or directs its attention towards) advertising messages to a very specific group of people

Extensive   large in scale, in-depth and detailed

Recap   repeat the main points

Misleading   giving the wrong idea or impression

Clarified   made them clear

Authorise   give permission to

Tend   usually do it

Juicier   causing more curiosity and interest

Nearest and dearest close friends and relatives

Worthless   having value or worth

Unusable   cannot be used

Third party   a group or organisation less directly involved

Manipulate   control in an unfair or dishonest way

We'd love to get your feedback on this episode.

What did you like? What could we do better?

What did you struggle to understand?

Let us know in the forum community.leonardoenglish.com