

École polytechnique de Louvain

Study of the influence of social networks on society

Author: **Jérémie KUPERBLUM**
Supervisor: **Peter VAN ROY**
Readers: **Olivier GOLETTI, Jérémy GROSMAN**
Academic year 2022–2023
Master [120] in Computer Science

Master thesis

Master in computer science

Specializations: Artificial Intelligence & Networks

Study of the influence of social networks on society

By

Jérémie Kuperblum

on June 4, 2023

at the UCLouvain

University supervisor: Peter Van Roy

Abstract

In 2021, people spent an average of 4 hours per day on their smartphone. When examining the most popular applications, social media platforms like Facebook & Instagram (Meta), TikTok, Twitter and Youtube (Google), emerge as the dominant players in the market. It is evident that there is a monopolization of this market, particularly by Meta and Google.

The business model of social media platforms is built on the principle of surveillance capitalism and the exploitation of human psychological vulnerabilities, such as the need for social validation and the desire to belong to a community. Their objective is to understand the deepest and most intimate secrets of their users, their sources of frustration, what makes them happy or unhappy, and more, with the intention of manipulating them in a progressive and imperceptible manner through persuasive technologies. Indeed, a study conducted by Facebook in 2012 demonstrated their ability to influence the mood of their users simply by filtering the type of posts shown on their news feed. Additionally, internal reports from Meta and TikTok have revealed that by measuring the time users spend on each post, they can determine their emotional states such as stressed, defeated, overwhelmed, anxious, and so on.

Another objective of social media have is to make their users addicted. In fact, everything is designed to give you a little dose of dopamine every time. This has several consequences, including making us addicted and more aggressive, reduced attention span, diminished cognitive abilities, decreased overall happiness, and more. It allows these platforms to condition and manipulate us through a technique known as intermittent positive reinforcement. Several former executives who have worked at companies like Meta and Google have openly admitted that the intention to make people addicted was intentional, and they themselves became addicted to their own platforms.

Furthermore, applications such as Instagram and TikTok have contributed to an increase in anxiety, depression, loss of self-confidence, and suicidal ideation among adolescents, particu-

larly young girls. Several internal studies conducted at Instagram have concluded that the platform is harmful. However, they do not address these issues as it would make their users less addicted. Instead, the platforms deliberately make users more vulnerable by further increasing their addiction, with the intention of manipulating them even further.

Facebook, Twitter, Instagram, and other platforms have also been utilized by entities such as Cambridge Analytica, the Internet Research Agency (IRA), various politicians, and governments to manipulate people, incite hatred, and propagate propaganda. Cambridge Analytica managed to obtain the Facebook data of 80 million users to assist several American politicians, including Mr. Donald Trump, in their presidential campaigns. They also played a role in influencing the British population to vote for Brexit. The posts disseminated by the IRA, whose objective was to support Mr. Donald Trump, reached 126 million Facebook users, 32 million Twitter users, and garnered 185 million likes on Instagram between June 2015 and August 2017. Additionally, Facebook is currently being used for activities such as human and child trafficking, hiring hitmen, suppressing press freedom in Vietnam, and so on.

We have observed that in several countries, including India, Myanmar, and Ethiopia, Facebook lacks an adequate number of moderators, and their artificial intelligence is incapable of understanding local languages and dialects. Furthermore, Meta often detects situations that violate their usage policies but does not resolve these issues until they gain media attention. Additionally, Meta has a program called XCheck that includes the names of personalities, politicians, and journalists who can violate their usage policies without facing consequences.

Social media platforms contribute to the creation of echo chambers due to filter bubbles. Combined with the fact that anyone can post whatever they want, that social media prioritize sensational posts because they generate more reactions, and that individuals tend to believe articles that align with their own ideas and surround themselves with like-minded people, this promotes the spread of fake news. In fact, it takes six times longer for true information to reach 1,500 people compared to fake news. Furthermore, social media has contributed to political polarization worldwide by trapping individuals in echo chambers. Consequently, people become less open to discussion and more aggressive towards those with differing political ideas.

We conducted four experiments on Facebook to investigate whether it confines users within echo chambers. Throughout these experiments, we encountered various challenges, including account suspensions. As a result, we were unable to fully complete our experiments. However,

the reasons behind these challenges are themselves worthy of study. Indeed, social media platforms take advantage of information asymmetry (they know everything about us, while we know nothing about them) to protect themselves. In particular, the fact that their databases are private, even for researchers, hinders the study of their platforms.

We have concluded that social media platforms are one of the causes of the increase in political polarization worldwide and have harmful effects on individuals. The fact that they have become one of the main sources of information for many people, that we are trapped in echo chambers, that their platforms amplify conflicts and prioritize sensational content, etc., will deteriorate our democracies if we do not take action. Furthermore, the world is now under the control of artificial intelligence (AI). Indeed, the recommendation algorithms primarily rely on AI that analyzes all our actions to provide us with content. It is therefore the AI that decides what the world should know by promoting certain information and preventing us from accessing others. Ultimately, the solution to all these problems can only come from us. It is necessary to raise awareness among people about the consequences of social media and all the concepts associated with it (echo chambers, filter bubbles, etc.). The most effective solution would be to not use social networks.

Acknowledement

Firstly, I would like to thank my supervisor Peter Van Roy who agreed to supervise me for this thesis, for all the precious advice, for the continuous support throughout the academic year and for all the time spent on meetings.

I would also like to thank Jérémy Grosman for taking the time to discuss my thesis with me and for the advice and various documents he gave me.

I would also like to thank Charles Pence for the various meetings, the documents he advised me to read and his availability to answer my questions.

Finally, I would like to thank Olivier Goletti and Jérémy Grosman for accepting to read my work.

Table of Contents

1. Introduction	1
1.1. Shaping our society in the Internet age	1
1.2. Research questions and objectives	2
1.3. Research Contributions	4
1.4. Structure of the Thesis	7
2. Experiments	9
2.1. First idea - September/October/November 2022	9
2.1.1. How to scrape Facebook ?	11
2.1.2. Difficulties	12
2.2. Second idea - November/December/January 2022	13
2.2.1. Difficulties	14
2.3. Third idea - January/February 2022	14
2.4. Fourth idea - February/April 2022	15
2.4.1. Scraping techniques	17
2.4.2. Difficulties	18
2.5. Untested experiment	19
3. The problem of the business model	20
3.1. Surveillance capitalism	20
3.2. When data collection is not enough	22
3.2.1. The manipulation of current data	22
3.2.2. Generate new data based on experiments	23
3.3. Social networks are a drug	26
3.3.1. How human nature drives us to spend more time on social networks . .	26
3.3.2. How social networks create the illusion of being alive	27

3.3.3.	Every action is designed to give you a shot of dopamine	29
4.	When the system is beneficial to third parties	33
4.1.	Advertising, the main source of revenue	33
4.2.	Cambridge Analytica	34
4.2.1.	Cambridge Analytica’s role in Trump’s 2016 electoral victory	37
4.2.2.	Cambridge Analytica’s role in influencing the Brexit vote	37
4.3.	Russia	38
4.3.1.	The IRA: an enterprise manipulating crowds on social networks	38
4.3.2.	How the IRA influenced the 2016 presidential election in favor of Donald Trump	39
4.4.	Sophie Zhang : a whistleblower’s account about fake engagement in politics . .	43
4.4.1.	Examining Facebook’s inactive response to election manipulation in Honduras	43
4.4.1.1.	When Facebook thinks it’s not important	44
4.4.2.	Facebook’s prioritization of Coordinated Inauthentic Behavior (CIB) .	45
4.4.3.	Powerlessness of Facebook in Azerbaijan due to a lack of moderators .	45
4.5.	Facebook’s amplification of hate speech and disinformation in Myanmar	47
4.5.1.	Facebook prioritizes its economic growth at the expense of its responsibilities	48
4.6.	Facebook, a platform for hate and violence in Ethiopia	50
4.7.	Facebook’s ad targeting tool strongly utilized in Europe	51
4.7.1.	Extremist political parties, the biggest spenders in social media advertising in Belgium	51
4.7.2.	Extremist Parties seem to invest heavily in Facebook advertising across Europe	52
4.8.	Non-exhaustive list of instances of negative impact by social networks	53
5.	Fake news	56
5.1.	Difficulties of the study of fake news	56
5.1.1.	Social networks: a platform conducive to fakes news	57
5.1.1.1.	The cost of disseminating information has become free	57
5.1.1.2.	The creation of clusters and echo chambers	57
5.1.1.3.	Social media platforms promote sensational and viral content	59

5.2.	Propagation of fakes news on social network during 2016 U.S. elections	61
6.	Possible solutions	65
6.1.	Regulatory measures: protecting citizens' privacy and financial penalties for social media platforms	65
6.2.	Investing in moderation tools and changing the business model	66
6.3.	Educating people on the impact of social media	67
6.4.	Stop using social media	67
6.5.	Individuals in key positions within social media companies do not use social media	68
7.	Conclusion	69
7.1.	The dark side of social networks: manipulation and control	69
7.2.	The information asymmetry and lack of transparency of social networks	69
7.3.	Social networks facilitates propaganda, manipulation, and human trafficking	70
7.4.	Social networks platforms prioritize fake news over authentic content	71
7.5.	The inaction of social networks in addressing these issues	71
7.6.	AI guides us, the world is under its control	72
7.7.	Social networks are at the origin of the rise of political polarization	72
7.8.	Future challenges: AI and Deepfakes	73
7.9.	The potential solutions	73
7.10.	Final words	74
List of Figures		75
Bibliography		81
A.	Source code of the experiments	123
A.1.	Codes for the first experiment	123
A.1.1.	Sources used that help us for the code	124
A.1.2.	Module.py	124
A.1.3.	Bot_test.py	125
A.1.4.	NewBot.py	129
A.1.5.	Script_Navigation_bot.py	156

B.2.	How to become a ghost on the web?	212
B.2.1.	The usefulness of a VPN	213
B.2.1.1.	Won't I be suspicious by using a VPN for Facebook ?	213
B.2.1.2.	VPN that we will use	214
B.2.2.	Browser fingerprint	214
B.2.2.1.	What is a browser fingerprint ?	215
B.2.2.2.	How it works ?	215
B.2.2.3.	Is it really used by websites ?	218
B.2.2.4.	Yes but my fingerprint can change over time according to my use of my computer. So I'm not longer traceable ?	218
B.2.2.5.	Why browsers allow web site to get those information ?	220
B.3.	American supermarket chain Target	221
B.4.	Experiments conducted by Mr. B.J. Fogg	222
B.5.	Cambridge Analytica	223
B.5.1.	The story behind the creation of Cambridge Analytica	223
B.5.2.	Brexit	225
B.5.3.	Russia	226
B.5.4.	Structure of SCL & Cambridge Analytica Companies & Shareholders .	227
B.5.5.	Cambridge Analytica - graphs	228
B.5.6.	Cambridge Analytica - Brexit	230
B.5.7.	Cambridge Analytica - Example of advertisements	231
B.5.7.1.	Example 1	231
B.5.7.2.	Example 2	231
B.5.7.3.	Example 3	232
B.5.7.4.	Example 4	232
B.5.7.5.	Example 5	233
B.6.	IRA	233
B.6.1.	Example of advertisements and posts made by the IRA	233
B.6.1.1.	Example 1	233
B.6.1.2.	Example 2	234
B.6.1.3.	Example 3	236
B.6.1.4.	Example 4	236
B.6.1.5.	Example 5	237

B.6.1.6. Example 6	237
B.6.1.7. Example 7	238
B.6.1.8. Example 8	238
B.6.1.9. Example 9	239
B.6.1.10. Other examples	239
B.6.2. The Top 20 IRA Facebook Pages	239
B.7. Europe and Facebook advertising	240
B.7.0.1. Hungary	241
B.7.0.2. Italy	241
B.8. Where do people get their news?	242
B.9. We are impatient	256
B.10.A society more divided than ever	257
B.10.1. Democrat and Republican, more divided than ever	257
B.10.2. What kind of world do we live in?	260
B.11.Study of the mentality of companies like Meta	262
B.11.1. Names of the speakers in the Netflix documentary The social dilemma	262
B.11.2. Topics covered in the documentary	263
B.11.3. Are we the product ?	264
B.11.4. Surveillance capitalism	264
B.11.5. Tracking	264
B.11.6. Do they sell our data to third parties ?	265
B.11.7. A world where online communication prevails over all others	266
B.11.8. Magic and social networks	266
B.11.9. Persuasive technologies (manipulation and addiction)	266
B.11.10A department dedicated to persuasive technology	268
B.11.11Is social networking a drug ?	271
B.11.12IA and social networks	274
B.11.13The problem of news feeds on sites like Facebook	275
B.11.14Polarization of society and loss of objectivity	278
B.11.15The problem of recommendations	279
B.11.16Fake news	280
B.11.17Politics	281
B.11.18Distinguishing truth from falsehood	282

1. Introduction

1.1. Shaping our society in the Internet age

Our societies have undergone significant changes over the past twenty years. They have placed services created by private companies at the core of their communication, production, consumption, and information flow systems. The advent of the Internet, high-speed connections, advancements in computing power, the decrease in data storage costs, the development of artificial intelligence, and so on, have completely revolutionized the way our societies operate. Everything has become digital. However, what are the implications of this transformation? After conducting some research, we observe that this system, on which our modern societies rely, is managed by a few companies that have capitalized on these technological advancements, including: (1) Google with its search engine, YouTube, and other services, and (2) Meta with its social networks: Facebook, Instagram, Messenger, and WhatsApp. If we also include Apple, Netflix, Amazon, and Microsoft (i.e., the GAFAM + Netflix), they account for 55% of the world's internet traffic [33].

As mentioned previously, our lifestyle has undergone significant changes, particularly with the advent of smartphones, which account for 65% of Internet traffic [34]. On average, individuals worldwide spent four hours a day on their smartphones in 2021 [73][135], and in France, 92% of teenagers between the ages of 12 and 17 own a smartphone [35]. So, what do we do on our smartphones that occupies so much of our time? We primarily scroll through posts and watching recommended videos. The most widely used applications are social media platforms such as Facebook, Instagram, TikTok, YouTube, and so on [211][212][210]. Similarly, among the most visited websites on the Internet, we find Facebook, YouTube, Twitter, etc. [112][11][112]. Additionally, Google Chrome is the most popular web browser with a market share of approximately 60-65% in 2022 [31][32]. The information provided here is a summary

of the content explained in Appendix B.1, where you can find more detailed research and graphs.

Given the prominent position they have taken in our lives and societies, it is important to understand the (negative and positive) impact they have on us and our societies.

1.2. Research questions and objectives

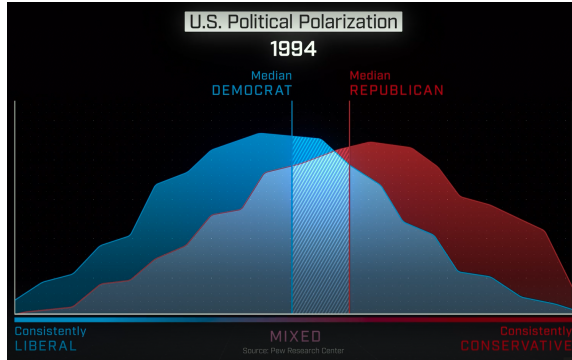
In this thesis, we will address several topics that are all intrinsically linked.

One notable observation we have made is that our world has never been as polarized, specifically within the political sphere. By analyzing the evolution of political polarization between Democrats and Republicans in the United States (figure 1.1), we observe that polarization remained relatively stable between 1994 and 2004. However, after 2004, polarization began to steadily increase. It is interesting to note that Facebook became accessible to the general public in 2006 [209], and social networks gained popularity on smartphones between 2008 and 2010 [106][179]. Furthermore, Facebook quickly emerged as the leading social network, amassing 500 million users by 2010 [154] (refer to Appendix B.1.8 for relevant graphs). So, there is clearly a correlation between the arrival and utilization of social networks and the increase in political polarization worldwide, but is there a causal link? This is one of the inquiries we aim to elucidate. This is one of the questions to which we will provide an answer.

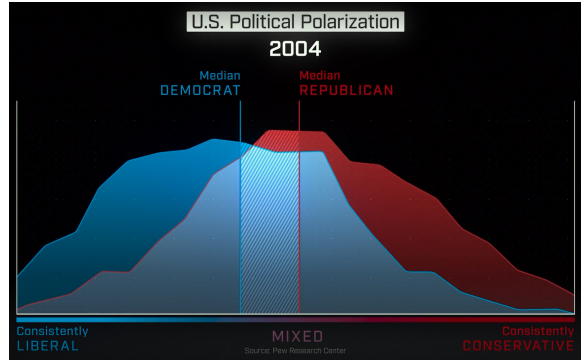
So, our first question pertains to the impact of social networks on societies and the world in general. What about their impact on individuals? Indeed, we have raised the question of whether these social networks have a negative and/or positive impact on the individuals who utilize them.

We have primarily focused our study on Facebook (and Meta more broadly) due to its status as the company with the largest user base in this market. However, this does not imply that our investigation is confined solely to them.

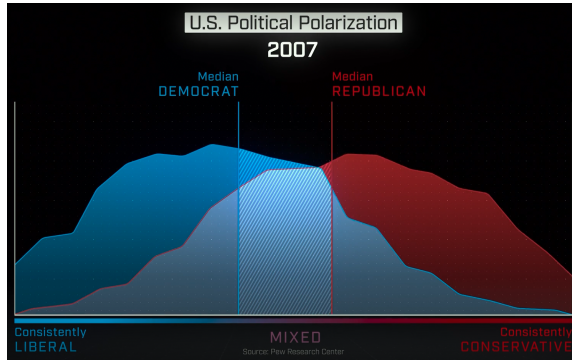
The graphs presented on the following page are from the Netflix documentary [264], which utilized data from the Pew Research Center. Additional graphs can be found in Appendix E.1.



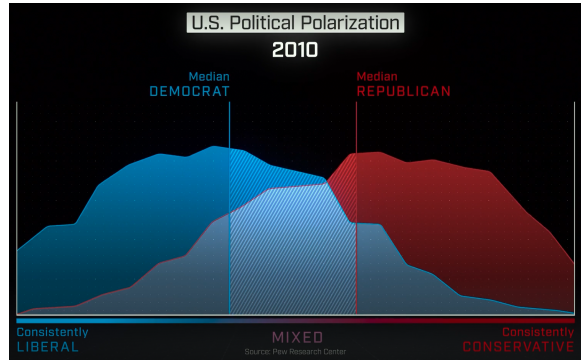
(a) U.S. Political Polarization in 1994



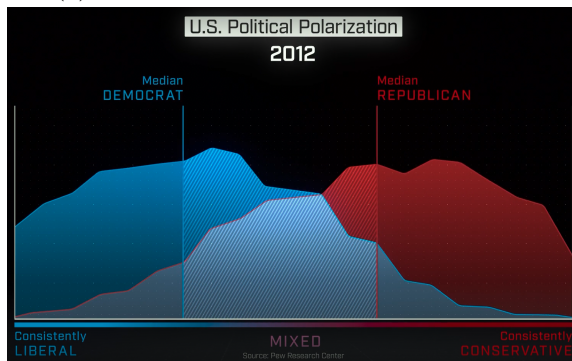
(b) U.S. Political Polarization in 2004



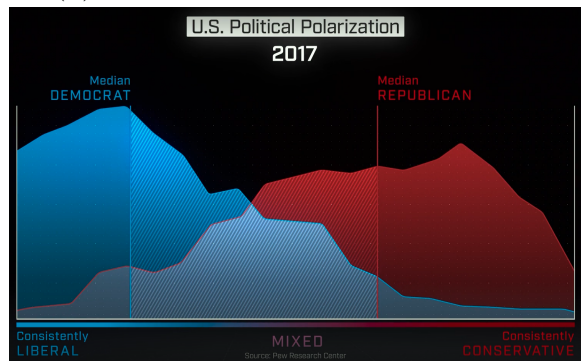
(c) U.S. Political Polarization in 2007



(d) U.S. Political Polarization in 2010



(e) U.S. Political Polarization in 2012



(f) U.S. Political Polarization in 2017

Figure 1.1.: Evolution of political polarization in the U.S. from 1994 to 2017. [264]

1.3. Research Contributions

- Facebook is a platform where conducting experiments to understand its functioning is challenging. Indeed, we encountered difficulties during our experiments, which prevented us from completing them. Social media networks refuse to share their databases with researchers or allow them to conduct experiments. They aim to maintain an asymmetry of information (they know everything about us, while we know nothing about them) as it protects them and grants them power.
- Their business model is based on surveillance capitalism. Their objective is to know every detail of our lives, including our most intimate secrets. Simply by measuring the time we spend on each post, they are capable of understanding our emotions and manipulating us.
- One of their other objectives is to keep us on their platform for as long as possible and to make us come back as frequently as possible. For this purpose, everything has been designed to provide us with a dose of dopamine every time. This has the consequence of making us addicted, reducing our cognitive abilities such as concentration, and making us more aggressive.
- In addition to making us addicted, this dopamine serves a second role, which is to manipulate us. Indeed, it allows social media platforms to do what we call intermittent positive reinforcement. It is a procedure where the likelihood of a behavior occurring increases when a stimulus is added in response to that behavior. By giving us a small dose of dopamine with each of our actions on their platform, they condition us to perform them until they become automatic.
- In 2012, Facebook conducted an experiment on 689,000 individuals without their knowledge. The objective of the experiment was to investigate whether the content presented to individuals on their news feed, including posts, comments, images, and more, could influence their emotions. The findings of this experiment revealed that *when the display of positive expressions was reduced, individuals generated fewer positive messages and more negative messages. Conversely, when the display of negative expressions was reduced, the opposite pattern emerged*. This demonstrates that Facebook has the ability to influence the emotions of its users based on what it shows them.

- Many people working in companies such as Meta, Google, Twitter, etc. have received training in persuasive technology. These are technologies whose purpose is to modify people’s behavior. Persuasive techniques are largely based on the theories of Dr. Brian Jeffrey Fogg, a social scientist and professor at Stanford University.
- Visual content is processed 60,000 times faster than text by our brains. It can easily evoke emotions and memories. This is the foundation of platforms like Instagram and TikTok. Combined with recommendation algorithms the fact that people seek social validation on these platforms, they are the cause of the increased anxiety, depression, and suicidal ideation among young people, particularly among young girls.
- Several internal studies conducted at Instagram have shown that the platform is harmful to its users. However, no measures have been put in place to address this issue because it suits them to have users who are mentally vulnerable and therefore more manipulable.
- TikTok’s algorithm had shown videos containing sex and drugs to accounts registered as minors
- Social media platforms are often used by governments, companies, and politicians to propagate propaganda, manipulate people, and fuel hatred towards other peoples:
 - Cambridge Analytica succeeded in obtaining the Facebook data of 80 million users to assist Mr. Donald Trump in his presidential campaign in 2016 and to influence the British public to vote in favor of Brexit.
 - During the 2018 elections in Brazil, a significant amount of fake news circulated on WhatsApp in favor of Mr. Jair Bolsonaro.
 - The IRA (Internet Research Agency) also made posts to assist Mr. Donald Trump during his presidential campaign. They managed to reach 126 million users on Facebook, 32 million on Twitter, and received 185 million likes on Instagram.
 - In Myanmar, Ethiopia, and India, Facebook has been used by governments and groups of people to incite hatred towards other ethnic or religious communities within the countries. Ethiopia, this has resulted in the death of 600,000 Tigray people and 2.6 million Tigrayans have fled the country. In Myanmar, approximately 9,000 Rohingyas have been killed and 700,000 to 900,000 Rohingyas have fled the country.

- Facebook is used for human trafficking, training and paying hitmen, selling children for sexual exploitation, and other illicit activities.
 - The Vietnamese government has mandated Facebook to censor posts criticizing the government, threatening to block Facebook in the country if they do not comply. Facebook has agreed to censor these posts.
 - In Belgium, it is the extremist parties (N-VA, Vlaams Belang, and PTB+PVDA) that invest the most money in advertising on Facebook. Together, they account for 78% of the total expenditure, across all parties, on Facebook advertising, which amounts to around 3.8 million euros.
- Facebook moderators do not understand all languages and dialects in every country, and Facebook does not have moderators in every country. Furthermore, their AI system for detecting content that violates their Terms of Use, does not comprehend all languages and dialects. For instance, it was only in 2018 that their AI began to understand Hindi (the most widely spoken language in India), despite India being the country with the largest number of Facebook users.
 - Meta detects instances of propaganda, hate speech, and the creation of groups targeting other people. However, they do not address these issues until they receive public and media attention. Moreover, when employees at Meta discover cases of manipulation, the executives are reluctant to allocate the necessary resources and time to address them.
 - Facebook has a program called XCheck that includes a set of politicians, celebrities, and journalists who can violate their Terms and Conditions of Use without facing consequences.
 - Recommendation algorithms create a bubble filter. This means that they filter the content that is presented to us based on our data (interests, browsing habits, etc.). This also contributes to intellectual isolation of individuals. In fact, everyone receives a different personalization of contents. As a result, we all see a different state of the Internet.
 - Recommendation algorithms also contribute to trapping us into echo chambers. Due to the fact that they only recommend posts aligned with our interests, we no longer receive information that go against our ideas. This has the consequence that individuals become less open to debate and there is an increase in hatred of opposing ideologies.

- Social media platforms contribute to the spread of fake news. They prioritize sensational and viral posts, even if they are fake news, at the expense of accurate and less sensational information. Sensational posts generate more reactions from users, and they spend more time engaging with them, allowing social media platforms to generate more revenue. Moreover, Mr. Zuckerberg has often put forward the advantage that virality had for the economy of his company which relies on their ability to encourage users to come back to the platform. One piece of evidence is that it takes six times longer for true information to reach 1,500 people compared to fake news on Twitter. When employees at Facebook propose moderating the sharing feature, the executives strongly oppose because it would impact the company's revenue. As stated by John Hegeman, Vice President and Head of Monetization at Meta, *If we remove a small percentage of reshares from people's inventory, they decide to come back to Facebook less.*
- The world is now under the control of artificial intelligence (AI). Indeed, the recommendation algorithms primarily rely on AI that analyzes all our actions to provide us with content. It is therefore the AI that decides what the world should know by promoting certain information and preventing us from accessing others.
- The solution to all these problems can only come from us. It is necessary to raise awareness among people about the consequences of social media and all the concepts associated with it (echo chambers, filter bubbles, etc.). The most effective solution would be to stop using them.

1.4. Structure of the Thesis

In the next chapter 2, we begin by presenting all our experiments, how we implemented them, and the difficulties we encountered. This chapter will focus on the negative impact of social networks on individuals. All the chapters following the experimentation chapter are based on literature research. We will start by explaining the problem related to their business model in chapter 3. This chapter will focus on the negative impact of social networks on individuals. We will explore the concept of surveillance capitalism, how they generate data about us, how they manipulate us with dopamine, and the fact that despite being aware of the harmful effects of their platforms, they do not address these issues due to their economic interests.

In chapter 4, we will examine how malicious third parties take advantage of social networks. We will specifically discuss Cambridge Analytica, the IRA, human trafficking on the platform, and the role of Facebook in contributing to the amplification of violence in India, Ethiopia, and Myanmar, etc. In addition, Meta does not invest enough in the effective management of its platform. We will continue our discussion in chapter 5 on the topic of the spread of fake news, where we will explore how they are prioritized over real information by recommendation algorithms due to their sensational nature, which often generates more reactions. The chapter 6 will examine several potential solutions to address these issues. Lastly, we will conclude our discussion in chapter 7.

A significant portion of the content has been placed in appendices to include only the essential information in the main text. When we refer to the appendices, please don't hesitate to consult them. Additionally, there are other appendices that are not directly referenced in the main text but are interesting to consult, such as Appendix B.11. We encourage you to explore them at your convenience.

2. Experiments

In this chapter, we will explain the experiments we conducted. We encountered several difficulties that prevented us from fully completing our experiments. Some of them would have been feasible if we had a larger team working full-time on this project and more resources. Moreover, this allowed us to realize that conducting a study on Facebook is very challenging, as they have implemented numerous tools to prevent research from being conducted. However, these experiments can provide ideas to other people and can be conducted again with more resources. Nevertheless, we were able to draw valuable conclusions from these experiments. In the last section 2.5, we will present an experiment that we would have liked to undertake if we had the time and resources.

The main objective of our experiments was to investigate whether the emergence of social networks is the causal factor behind the rise of political polarization worldwide. Recommendation algorithms are complex systems that rely on our actions, the time we spend on each post, whom we interact with, what we like, comment and share, etc. Everything is tracked to select the content that will be displayed to us. The question is therefore whether these algorithms lock us into our ideas by offering us only what aligns with our beliefs and create clusters of like-minded individuals (or not). If this is the case, it diminishes our receptiveness to alternative ideas and it can be the cause of the increase in political polarization worldwide.

2.1. First idea - September/October/November 2022

Following our research in September, we started to think about our first experiment. The objective of this experiment was to determine whether we would be confined within our ideology or if we would receive diversified content on our news feed. Another question we aimed to address was to determine the criteria upon which Facebook algorithms rely to propose posts

on your news feed. For example, will fake news that receives a high number of reactions (likes, comments, shares) be given greater visibility than true information that receives fewer reactions? In this case, the algorithm prioritizes the quantity of reactions over the accuracy of an article. This is what we intend to ascertain.

In order to do so, we came up with the idea of using the *web scraping* technique, which allows us to extract the content of the webpage we are on using a script (our scraping technique is explained in the subsection 2.1.1). The objective was to collect everything that Facebook offers us: all the posts and their type (article, video, picture, ...), the number of reactions (likes, comments, shares), the date when the post was published, the author, ... By collecting all this data, it would have allowed us to analyze them with another script in order to try to discover the parameters on which the recommendation algorithm is based to propose us content. For instance, if the algorithm detects that I have an interest in the stock market and that I follow the news page Le Monde, will it only recommend content related to the stock market? If that's the case, what proportion of the recommended content is dedicated to this topic? If Le Monde publishes 10 articles throughout the day, what subjects will Facebook propose to me on my news feed? Does the algorithm primarily prioritize the content's relevance to my interests or the number of reactions it receives? Which parameter holds greater significance for the algorithm? Moreover, we intended to examine the criteria Facebook relies on when filtering comments to suggest the most relevant ones. To achieve this, we would have collected all the comments filtered as relevant and compared them to the unfiltered comments. Small note: when collecting data, we would have anonymized the names of all non-public individuals.

Now that the general idea is outlined, how do we go about implementing it? We intended to create accounts with diverse personalities, varying political ideologies (communist, capitalist, etc.), different age groups, different genders, in various countries (using a VPN), individuals who believe in conspiracy theories, and so on. In order to encompass a wide range of ideologies and facilitate the formation of large clusters of interconnected accounts, we wanted to create a hundred accounts. We also wanted the accounts to not all be located in the same country in order to observe if there are differences between two accounts with the same ideology but located in different places. We had thus chosen to use a VPN, specifically NordVPN (refer to Appendix B.2.1 for an explanation of what a VPN is and its purpose).

So at the beginning of the experiment, all our accounts were ideologically neutral as they had not yet displayed any interest in a particular subject. Furthermore, they were not yet

connected to any other accounts and did not have any friends. It would have been intriguing to observe what Facebook proposes when it is unaware of our preferences. Does it prioritize certain topics? Does it exclusively suggest posts that have received significant engagement? And so on. Then, the accounts would have remained friendless but would have begun to exhibit a growing interest in a specific subject matter, such as capitalism, communism, conspiracy theories, and so on. To accomplish this, we would have followed pages, liked, commented, and shared posts, as well as shared articles from the internet on Facebook. Our aim would have been to observe the changes in their news feed. Would Facebook primarily suggest posts that align with their ideology, thereby limiting their exposure to a single topic and perspective? Or would Facebook strive to diversify the content, offering a broader range of topics and viewpoints?

In the third step, we would have interconnected certain accounts to form clusters of ideas. For instance, we would have created a cluster consisting of four accounts with a capitalist ideology, another cluster with two capitalist accounts and two communist accounts, and so on. We could have explored various combinations and observed how the news feeds of different cluster types evolve. For instance, if a person with a capitalist ideology becomes friends with a person with a communist ideology, would Facebook suggest posts from the communist friend on their news feed?

Finally, we would have created a large network where all accounts would be interconnected either directly or indirectly through a mutual friend. Thanks to this network, we would have been able to analyze, on one hand, the evolution of the news feed, and on the other hand, the dissemination of information within a vast network of individuals. For example, if person A (with capitalist ideas) shares an article, which other accounts would have encountered this article on their news feed? So, we would have been able to determine whether, in a large network of people, Facebook creates clusters in which only one type of content circulates. This would have allowed us to assess whether Facebook restricts open-mindedness by confining individuals within their own ideologies.

2.1.1. How to scrape Facebook ?

After conducting several days of research, we had decided to utilize the Selenium module with the Python language. This module enabled us to develop a bot that automates various tasks

for us, including browsing and logging into Facebook, extracting information from HTML pages (scraping), clicking on buttons, and more. The Selenium module utilizes the HTML code of the webpage it navigates to extract information. For instance, when the bot intends to log in to Facebook, it opens a web page, accesses Facebook, identifies the username and password fields, enters the login credentials, and clicks on the login button. All of these actions were performed by examining the HTML code to locate the elements. To develop our bot, we created an initial fake account which served as the basis for our coding.

Our bot was capable of logging into Facebook and accepting cookies, scrolling through the content, clicking on comments to expand them or change their display order (either by most recent or most relevant), clicking on post reactions to determine the quantity of each type (number of likes, number of loves, number of hates, etc.), retrieving information about the author, description, and the timestamp of the post, and so on. You can find the different codes for this experiment in the Appendix A.1.

However, we encountered several challenges that made the experiment more complex than initially anticipated, which we will explain you in the next subsection 2.1.2.

2.1.2. Difficulties

Another difficulty we encountered was the inconsistency in Facebook’s HTML structure, which posed challenges for the bot’s navigation and data extraction. Specifically, when trying to locate the author’s name for each post, the Xpath (the path in the HTML) varied and was not consistently the same. Hence, in the code, we implemented multiple Xpaths to handle different variations. However, there was still a risk of missing certain Xpaths and consequently not retrieving all the desired data. Another challenge we faced was our dependency on Facebook’s HTML structure, which is subject to frequent updates. Any modifications to the HTML would have necessitated recoding the bot by adjusting the Xpaths used for navigation and scraping. This would have posed significant maintenance and updating challenges.

Another challenge we encountered was the management of a hundred accounts simultaneously and the need to change locations for each account. However, due to time constraints and technical limitations, it was not feasible to manually handle this process, especially if we intended to establish multiple connections per day using a single computer.

Ultimately, the core challenge we encountered lay in the use of a bot, which rendered us vulnerable to detection by Facebook. The Selenium module itself is detectable by Facebook’s systems. Although there are solutions to avoid detection, this has added an additional level of complexity to our experiments [9]. Additionally, the browser (Firefox) detects when he is control by a script, and Facebook would have quickly realized the absence of human activity based on the bot’s behavior. For example, our bot did not use a mouse to navigate and Facebook uses mouse trackers [110]. If Facebook detects that it is a bot, it can either ban the account or put it in another Facebook environment where there are only bots [225]. This would have significantly compromised the validity of the experiment. In addition, there is another related problem, which is the browser fingerprint (we explain what it is in Appendix B.2.2). Indeed, despite our utilization of a VPN to relocate and alter the IP address for each account, all of these accounts were connected from the same computer. Therefore, if the IP address is not taken into account, all of the accounts would have had the same browser fingerprint, further raising suspicions from Facebook.

We persisted with this experiment for several weeks, exploring various solutions to our challenges. However, we came to the realization that it would not be feasible. Even if we had managed to collect data, the experiment would have been highly vulnerable to criticism.

2.2. Second idea - November/December/January 2022

For this experiment, we took inspiration from the previous one but redesigned it in terms of the number of accounts and techniques employed. Specifically, we aimed to limit ourselves to 5 accounts, which would have been more manageable for us to handle without relying on bots. Alternatively, we explored alternative approaches to the Selenium module and came up with the idea of creating a browser extension. This extension would allow us to capture the HTML of web pages simply by clicking on the extension button. Subsequently, using a Python script, we would have extracted the required information from the downloaded HTML files.

With this approach, we eliminated the use of an automated script (i.e., the bot) and performed all tasks manually. The advantage was that the browser extension we developed was undetectable by Facebook, or at least it would have gone unnoticed.

We developed the extension script and the Python script to convert the downloaded HTML

files into CSV format, with each line representing the information of a post. The codes for this experiment can be found in Appendix A.2. Despite our efforts, we encountered various difficulties.

2.2.1. Difficulties

Initially, as part of the experiments, we attempted to automate the creation of emails using Outlook. We developed a bot that would generate emails for us using a predefined list of names and surnames, and it would also generate passwords automatically. The initial challenge we encountered was the requirement of human presence for completing the captcha during the email creation process, which was manageable for a smaller number of accounts. Unfortunately, Outlook detected our bot and required multiple captchas for each email creation attempt, which made the process lengthy. Furthermore, it appeared that Outlook could distinguish the artificial nature of the accounts, compromising their authenticity. When we attempted to create fake Facebook accounts using these emails, the accounts were banned, whether they were created by a bot or manually. Appendix A.5.3 provides screenshots of the banned accounts. We created 35 email accounts, but they were unusable. Then, when we tried to create a Facebook account using a manually created email, we were banned again. We suspect that our computer was blacklisted, as we were unable to create any more accounts without being banned, which prevented the experiment from taking place. The code to create emails and Facebook accounts using a bot can be found in Appendix A.5.

Another obstacle we faced was our dependence on HTML and the need to identify different Xpath variations for each type of information we wanted to extract. Indeed, the Python script that collected all the information we needed in the downloaded HTML files was using Xpath. Furthermore, the HTML code obtained through the extension had inexplicably missing information, rendering it impossible to retrieve all the required data using our script. Regrettably, we were unable to find a solution to this persistent problem.

2.3. Third idea - January/February 2022

To overcome the obstacle of our computer being blocked by Facebook from creating accounts, we devised an alternative approach. We decided to use virtual machines (VMs) as a solution.

Using Oracle VirtualBox, we created five VMs, each running on an Ubuntu 22.04 ISO, with one VM dedicated to each account.

Once the VMs were set up, we manually created an email account on Outlook within one of the VMs. Subsequently, we attempted to create a Facebook account using this email address. However, we were directly banned. The act of creating an account via a VM raised suspicion for Facebook.

2.4. Fourth idea - February/April 2022

We were still determined to conduct the experiment, but it required the participation of external individuals who could use their own computers without VPNs or VMs to minimize suspicion. Initially, we considered the option of hiring a small team. Ultimately, we decided to create the project for the LINFO1115 course of Mr. Peter Van Roy, centered around the subject of this thesis. The project was scheduled to last for 4 weeks and involved 73 students. However, this project did not take place, and we will provide an explanation for this in the subsection. 2.4.2.

The concept was as follows. We conceived four types of accounts: capitalist accounts in support of nuclear power plants, capitalist accounts against nuclear power plants, communist accounts in support of nuclear power plants, and communist accounts against nuclear power plants. The objective was to have opposing ideologies. The project was planned to be conducted in groups of two, with each group assigned a specific account type. Within each group, both students would have the same type of account. As part of the project, they were required to create a new user profile on their computer to ensure that their personal activity would not interfere with the project. Subsequently, they were expected to create a Facebook account using their UCLouvain email address. Indeed, based on our previous experiment, creating a Facebook account using a freshly created email raised suspicion from Facebook, resulting in our accounts being banned.

Therefore, at the initial stage, for a period of 2-3 days, the students would have been required to scrape their news feed (the technique we had implemented for scraping will be explained in the subsection 2.4.1) without following any pages or displaying any particular interest in a specific topic. The objective was to assess whether Facebook was already guiding individuals

towards certain topics.

During the following three days, the students were supposed to start following five news media pages from a pre-selected list of 14 options (such as Le Figaro, Le Monde, La Libre, Charlie Hebdo, etc.) and continue scraping their news feeds. The objective was twofold: firstly, to observe the evolution of the news feed and secondly, to analyze the articles suggested by Facebook on their news feed. For example: if Le Figaro publishes 10 articles per day on its page and we only receive 3 of them on our news feed, on what criteria did Facebook choose these three in particular? Is it because they generate more reactions? Is it because of the topic of the article? etc.

The second week had an interesting aspect. Each student would have been required to send a friend request to a person who had created an account of the same type but who was not part of the same project group (we would have determined the specific friend request connections). In addition, they would have allowed the Facebook algorithm to understand their ideology by following more specific pages from a list we created, by sharing articles (from external websites) on their Facebook profiles, and by engaging with articles through comments, likes, and shares on Facebook. The objective would have been to enable the algorithm to understand the ideology of their account. For example, if it was a capitalist account in favor of nuclear power plants, it would have followed/liked/shared/etc. numerous Facebook pages related to banks or investment advice. Additionally, it would have shared videos and articles from other websites explaining that nuclear energy is a good solution. Moreover, they would have had to send messages to their new friends and actively engage with their posts by liking them. The goal would have been to demonstrate a strong friendship and establish a small group of two individuals sharing the same ideas. Concurrently, they would have needed to continue collecting data from their news feed to observe its evolution. This would have allowed us to determine if they would have been trapped in their own ideology or if Facebook would have shown them some diversity in the types of articles suggested. Would they have continued to receive articles from sources like Le Figaro, Le Monde, and others?

The third week would have followed the same approach, but we intended to form groups of 4 individuals sharing the same ideology (ensuring that no member of the cluster would be friends with their project partner). The objective would have been to further reinforce this ideological grouping. Participants would have continued actively sharing, liking, and commenting on articles in line with the ideology of their account, while maintaining strong

friendships with their Facebook friends on the platform. The aim would have been to observe how their news feed would have evolved, addressing the same questions as in the second week. Finally, in the last week, we would have connected two clusters, each consisting of four people, resulting in groups of eight. The difference from weeks 2 and 3 is that this time there would have been clusters with at least one opposing ideology (capitalist vs. communist and for nuclear power plants vs. against nuclear power plants). When the two clusters would have been connected, each person would have had seven friends. However, they would have had to continue maintaining friendships with people sharing the same ideology (from their cluster of four) in order to maintain strong ties. They should not have sent messages or engaged in liking/commenting/sharing posts from people in the opposing cluster. This step would have allowed us to observe how their news feed would have evolved if they became friends with individuals who did not share the same ideas and if they did not have strong friendships with them on the platform. Maybe at the beginning of the friendship (the first 1-2 days), they would have all received the posts of everyone, and after these 2 days, they would have only received the posts of their ideology because they would not have shown interest in the posts of people from the other cluster. This would have shown that we can easily get locked into our ideas. On the other hand, we could have had the scenario where it would have led to a diversification of the content proposed on their news feed. Several scenarios could have occurred. Of course, throughout this stage, students had to continue to scrape their news feed.

2.4.1. Scraping techniques

Our previous scraping techniques were not conclusive, especially since they were dependent on the computer used. We also found that the HTML structure of Facebook is not always consistent. Therefore, we developed an alternative method. We abandoned the use of browser extensions and adopted a technique based on analyzing browser queries. When you are on Facebook, Facebook initially sends you an HTML page containing only the first post. The subsequent posts are transmitted through JSON files, which are processed by JavaScript scripts to add them to the HTML page. By retrieving these JSON files, it is possible to obtain the posts (or at least a portion of them, as we will explain in detail the difficulties) that follow the first post. To accomplish this, the students would have needed to use the Firefox browser.

To retrieve the data, it was necessary to access the browser's inspector and go to the network tab. From there, the students would have selected only the HTML requests before going to Facebook. Next, they would have had to download the .HAR (HTTP Archive) file after scrolling on Facebook. This file contained a collection of information and queries that were needed for analysis. With another Python script (which can be found in Appendix A.4), the students would have been able to analyze these files and extract the relevant information. Therefore, throughout the project, all students would have been required to retrieve this .HAR file each time before closing Facebook.

2.4.2. Difficulties

The first difficulty we encountered was the fact that we came up with the idea to use the .HAR files a little too late, and the Python script was not yet perfected at the beginning of the project. We had to figure out how to search for the desired information within the .HAR file, and some of this information were difficult to find. Furthermore, we realized that, for unknown reasons, certain queries were incomplete. As a result, we would not have been able to retrieve the information from every post, which could have compromised the accuracy of the experimental results.

Another major difficulty lay in the inherent risk of the project. Students could be banned from Facebook at any time, which would have had significant consequences on their project grades. Some students were even banned shortly after creating their accounts. They had not followed the recommendations in the instructions, such as the prohibition of using virtual machines, VPNs, or newly created email addresses. However, this underscored the risky nature of the project. If we had a dedicated team working on this project instead of students, it would have been more feasible (in theory).

We are not the only ones who have encountered these difficulties that we have explained to you. In fact, the YouTuber Poisson Fécon wanted to conduct an experiment on the Tinder application by creating fake accounts. In his video, he explains the challenges he faced in setting up his project, as all of his accounts were banned. However, he was able to obtain some results thanks to the team and resources he had. [159]

2.5. Untested experiment

We wanted to attempt to disseminate misinformation on Facebook to assess the level of ease with which people can be manipulated. We wanted to ensure that our misinformation had no consequences on any individual or entity, hence the need for an innocent subject. Our idea was to make people believe that UFOs were flying over Louvain-la-Neuve. Through our research, we discovered that the manipulation techniques employed on Facebook involved, firstly, creating a false community that supported the subject to give the impression that there was already a base of believers, and secondly, having the information posted by multiple sources. Therefore, we would have had to create several Facebook groups with a fabricated community. Creating Facebook groups is not a complicated task, and for the fake community, we intended to employ the same technique as explained in the section on Honduras 4.4. This would have involved creating a few fictitious accounts (for example 10) and using these accounts to create multiple pages that would have followed our Facebook groups. With 10 fictitious accounts, we could have easily garnered a base of over 100 followers for each of our groups. Regarding the posts we would have made, we would have had to edit photos and videos to publish them using different accounts, thereby giving the impression that they originated from diverse sources. Using our fictitious accounts, we would have liked, shared, and engaged in fake conversations to support the photos and videos. It would have been easy to gauge people's intrigue in the subject by observing if real individuals began to follow and interact with our pages and content. In addition to that, we also had the idea of advertising on Facebook to attract people to our groups.

This experiment could have been a failure or, on the contrary, elicited reactions from real individuals. Would we have succeeded in making people believe that there were UFOs flying over Louvain-la-Neuve? If so, it would have demonstrated how easy it is to manipulate people and reach a wide audience on Facebook.

3. The problem of the business model

3.1. Surveillance capitalism

Services offered by companies like Google, Facebook, and others are *free* for users. You can search on Google Chrome and scroll on Facebook without paying any fees. However, it costs these companies money to enable you to perform these actions. But where does their money come from? Simply from advertisements. Their sales product, their ingenuity, is thus hidden under this concept of advertising. And what they sell to their advertiser is **certainty** [Shoshana Zuboff in documentary The Social Dilemma [264]] [321] : when an advertiser pays companies like Google and Facebook to publish their ads, they guarantee that it will be successful! So far there is nothing wrong: they need to make a profit, and they generate it with advertising. But what lies behind this guarantee? That's what we are going to talk about now.

As an advertiser, your goal is to choose the platform that will allow you to target people most effectively with your advertising. It is in this market that platforms such as Facebook, Google, etc. compete to create the most precise targeting tools. For Facebook, the more money their advertisers make, the more money they will make because advertisers will continue to invest in advertising [264]. So, their business model relies on statistical models to predict and maximize engagement with content [238]. To make predictions, two essential elements are required: knowing how to collect relevant data and correctly analyzing/manipulating them [264][321].

So, to make predictions, it is necessary to understand the interests of users. This is not very complicated for platforms like Facebook, where they simply need to use trackers to collect data on what their users watch, the pages they subscribe to, and so on. Except that, other companies like Google can easily do the same thing by analyzing what their users search on

the Internet or watch on YouTube. It is important to remember that these companies are in competition, and their market is centered around effectively targeting their advertisers' ads in the most precise way [357]. Being satisfied with interests alone is too simplistic. That's why they entered the human behavior market [356]. Their goal is no longer to simply know our interests, but to understand their users perfectly: their relationships, their routines, what makes them happy or unhappy, what frustrates them, their character traits, whether they are extroverted or introverted, their political and economic ideas, their studies, their social status, and more. With all this data, they are able to perfectly control when and which advertisements to display on their users' news feed. But they go further than that: they attempt to modify our behavior [264][356][321] in their own interest, which can be, for example, economic [357]. All this data they collect is no longer used solely for the purpose of accurately targeting advertising, but also to **gradually** control their users in a way that is **imperceptible** from our perspective. It is well-known that the greater our knowledge about an individual, the greater our ability to control/manipulate them [357]. They no longer strive to predict our future; instead, they dictate it to us. They impose behavioral norms, manipulate our emotions by amplifying, diminishing, or imposing them [264]. They exploit our emotions [287] because it allows them to discover who we are and manipulate us (we will come back to this in section 3.3 where we will talk about dopamine). They impose upon us what they deem as good or bad for society, they lock us into a certain ideology through bubble filters (we will discuss this concept later). They have the ability to control everything due to their access to a vast amount of data, their storage and processing infrastructure, and their monopolization of networks. This enables them to analyze every aspect of our interactions with people (e.g., Meta with Facebook, Instagram, Messenger, and WhatsApp) and control the flow of information. This manipulation remains **imperceptible** due to two factors:

- There is an **asymmetry of information** (they know everything about us and we know nothing about them) and they have great control over the physical infrastructure [357]. They have the ability to collect any information they want about us, to decide what they will show us or propose to us, to guide us towards what they desire, and so on. They know, they decide and they decide who decides without us knowing it.
- They do it in a **progressive** and **imperceptible** way [Jaron Lanier [264]]. It's similar to societal norms; they evolve over time. However, we often don't realize when norms have changed because they occur gradually. It's only afterward that we become aware

of the changes. The fact that the change is progressive over time makes it imperceptible in the moment, and the underlying cause of the change becomes unidentifiable.

Everything we have just explained has a name proposed in 2014 by the American economist Shoshana Zuboff, who is a professor emeritus at Harvard Business School [322]: **the surveillance capitalism**. We will explore this concept in more detail in the following section.

3.2. When data collection is not enough

As mentioned before, their business model is based on the data of their users, and this data must be abundant and diverse [321].

They have direct access to a set of data through their platform and our usage of our devices: our friends, our browsing habits, our interests, the time we spend on each post, whether we are at home or in the office, and more.¹

Their objective is no longer simply to know how we behave on their platform. They seek to know the smallest characteristics of our lives, all our personality traits (even those we are not aware of), our lies, our weaknesses, and more. They want to capture and retrieve everything in order to create **certainty** and **manipulate** us, turning us into puppets whose strings they pull. What they aim to discover is not just our interests, but to understand who we are in every detail, gradually guiding us towards their own interests. To achieve this, they need to generate new data. They can do it in two ways:

1. The manipulation of existing data in order to predict/generate new ones [321].
2. Via experiments [321]

3.2.1. The manipulation of current data

When the collected data becomes insufficient in quantity, one of the solutions is to generate new data based on existing ones, such as deducing new data or making predictions [321]. Indeed, it

¹You can observe the data they are able to collect by visiting Meta's website: **Data for good**. However, it's important to note that the data available on this website is only a fraction of what they actually collect. They selectively present certain data while keeping the majority hidden. As mentioned earlier, there exists an asymmetry of information.

is no longer a matter of knowing what you currently consume, but rather of anticipating what you will consume. Numerous companies adopt this strategy driven by economic motives. For example, in 2002, the American supermarket chain Target wanted to predict if a woman was pregnant in order to send out discount vouchers on baby products before their competitors. Unfortunately, they succeeded. In 2012, a high school girl received coupons for baby clothes and cribs. That's how her family found out, which is catastrophic in terms of our right to privacy... [107][226] (more information in Appendix B.3). As for social media, it makes sense that they engage in similar practices, but with a higher level of precision.

3.2.2. Generate new data based on experiments

Once we have deduced and predicted the maximum amount of data based on the information we possess, why not generate new data through experiments? One might think that we would become aware of such experiments if they were being conducted. However, these experiments are **imperceptible**, which grants them immense power.

To support this point, we can mention the experiment conducted by Facebook in 2012 on 689,000 individuals without their knowledge [236][174][156]. The objective was to examine whether the content presented to individuals on their news feed (posts, comments, images, etc.) could influence their emotions. Specifically, they aimed to test the hypothesis of emotional contagion, which suggests that exposure to certain emotions can impact an individual's own emotional state. The findings of this experiment revealed that *when the display of positive expressions was reduced, individuals generated fewer positive messages and more negative messages. Conversely, when the display of negative expressions was reduced, the opposite pattern emerged*. These results suggest that the emotions expressed by others on Facebook have an influence on our own emotions, providing empirical evidence of widespread emotional contagion through social networks [236][227][174][178]. Facebook clearly became aware of the power it has by simply curating the content presented to us on our news feed.

When Facebook published the study, it sparked controversy [296][27][156]. Indeed, nobody was aware that an experiment was taking place on the platform (it went **imperceptible**). Moreover, it revealed the significant power that Facebook possesses in terms of manipulation, even as early as 2012. Simply by filtering the news feeds of its users, Facebook managed to exert an emotional influence on them. This one-week experiment yielded significant results...

[156][227]. Imagine what they have been able to implement since then. Due to the negative reactions generated by this experiment, Facebook has never again published a report on their experiments and has never mentioned conducting any new ones. However, that does not mean they have stopped conducting them!

Facebook has therefore understood that it had the opportunity to exploit human vulnerabilities to generate new data and manipulate its users according to its wishes. Sean Parker himself explains that the functioning of Facebook is based on the **exploitation of human vulnerability and psychology** [72]. Indeed, an internal Facebook report, leaked in 2017, revealed that the platform had the capability to discern when young people feel stressed, defeated, overwhelmed, anxious, nervous, stupid, silly, useless, and a failure simply by monitoring the posts, photos, and videos presented to them in real time [307]. Similarly, The Wall Street Journal found that TikTok only needs one piece of information to understand what a user is interested in: the duration of time they spend on each post [295][136]. In the source [295], they provide an example of how this mechanism operates, and in Appendix C.2, we explain how Instagram can identify your emotions and what triggers them. Moreover, on Instagram and TikTok, everything is visual, and humans strongly respond to visual stimuli. In fact, visual content is processed by our brain 60,000 times faster than text [94]. Through the use of visual content, memories and emotions can be triggered rapidly [245][327]. By discovering how you feel and what triggers these emotions, they can uncover the most intimate aspects of your life. With the information they have discovered, they can keep you engaged on their platform by manipulating your emotions.

Meta has conducted multiple studies to comprehend the impact of Instagram on its users. Each study has concluded that Instagram is **harmful**. Indeed, Instagram contributes to the increase in anxiety, depression, body dissatisfaction, eating disorders, and some young individuals attribute their suicidal thoughts to Instagram. These negative effects are particularly prevalent among girls [169]. This is partly due to the fact that we experience emotional fluctuations when using these applications, by constantly comparing ourselves to others and seeking social validation. However, this does not mean that these platforms will address these problems as it serves their interests. What is most astonishing is that Meta is aware of the detrimental impact Instagram has on the well-being of young people, yet it publicly claims the opposite and makes minimal efforts to address these issues. When questioned about the findings of their internal research, they refused to disclose them [170]. Once again, we are confronted

with an asymmetry of information that is advantageous to them. Feel free to consult these sources from The Wall Street Journal: [170][169], which provide a wealth of interesting details and served as the basis for writing this paragraph.

The youth market (as we have seen with Instagram) is an audience that interests social media platforms. This is due to the ease with which they can be manipulated and their immersion in technology since birth. Instagram, TikTok, Snapchat, and YouTube are competing to attract this audience. Meta has notably conducted a dozen studies to understand which products could be popular among preteens (aged 10 to 12) [169]. Preteens are no longer protected from advertising (for instance, on television, it is prohibited to broadcast advertisements during children's programs [352][157]) and can be too easily exposed to inappropriate content. As an example, The Wall Street Journal showed that TikTok's algorithm had shown videos containing sex and drugs to accounts registered as minors [295]. [169]

Finally, Instagram is the perfect app to get you hooked. It primarily consists of images and videos, has a predominantly young user base (40% are under 23 years old [170]) who are highly susceptible to manipulation, and relies principally on content recommendations.

So, social networks are not tools that wait to be used, they have their own objectives, and to achieve them, they employ psychological techniques against us [Tristan Harris² [264]]. They employ artificial intelligence to comprehend the workings of the human brain, in order to uncover what triggers emotional responses [264]. They are well aware of the flaws in the human brain and how to exploit them. To prove this, Sean Parker confessed it during an interview with the Axios newspaper [257]. He explained that: *When Facebook was getting going, I had these people who would come up to me and they would say, 'I'm not on social media.' And I would say, 'OK. You know, you will be.' And then they would say, 'No, no, no. I value my real-life interactions. I value the moment. I value presence. I value intimacy.' And I would say, ... 'We'll get you eventually.'* This already demonstrates their objective to succeed in *getting us*. We can sense in what he says that he seems to know things we are unaware of. He also displays a sense of superiority towards people. However, he later admitted that he did not realize the consequences of what he had said: *The thought process that went into building these applications, Facebook being the first of them, ... was all about: 'How do we consume as much of your time and conscious attention as possible?' And that means that*

²Tristan Harris is the Co-Founder & Executive Director of the Center for Humane Technology (CHT) [Information found on his website: <https://www.tristanharris.com/>]

we need to sort of give you a little dopamine hit every once in a while, because someone liked or commented on a photo or a post or whatever. And that's going to get you to contribute more content, and that's going to get you ... more likes and comments. It's a social-validation feedback loop ... exactly the kind of thing that a hacker like myself would come up with, because you're exploiting a vulnerability in human psychology. The inventors, creators it's me, it's Mark [Zuckerberg], it's Kevin Systrom on Instagram, it's all of these people understood this consciously. And we did it anyway. [257][279][311].

To support the fact that all of this was done consciously, Chamath Palihapitiya³ clearly explains that the goal of these companies is to manipulate us psychologically as much as possible by providing us with a dose of dopamine with each of our actions (known as positive reinforcement, explained in more detail in section 3.3) [264][299][220]. He explains that the goals he was given were for every new user to have at least 10 friends after the first week of joining the platform. Sandy Parakilas explains that Chamath Palihapitiya employed techniques where he conducted several tests on the Facebook platform and observed user reactions. Through these tests, they were able to optimally determine what users expected from their services [264]. So, it is manipulation, and Chamath Palihapitiya himself states, *You don't realize it, but you are being programmed* [299]. The question is: how do they succeed in providing us with this dose of dopamine? We will address this question in the next section 3.3.

3.3. Social networks are a drug

3.3.1. How human nature drives us to spend more time on social networks

Social networks exploit our instinctive behaviors. Indeed, Dr. Anna Lembke⁴ explains that it is encoded in our genes that we need to live in communities in order to reproduce and perpetuate our species. We have a natural need to be connected to others, to be together [264]. We instinctively react to our environment, to what we perceive as being alive [161], and we are naturally drawn to what brings us together. Social networks provide us with this sense

³Chamath Palihapitiya was former vice-president of user growth from 2007 to 2011 at Facebook. His LinkedIn: <https://www.linkedin.com/in/chamath/>.

⁴Dr. Anna Lembke is a professor of psychiatry at Stanford University School of Medicine and chief of the Stanford Addiction Medicine Dual Diagnosis Clinic [76].

of connection, of belonging to a community. A Harvard study and other studies have shown a strong correlation between increased happiness in an individual and the presence of strong social relationships [252][330]. Why? Personal connection creates mental and emotional stimulation, which are automatic mood boosters, while isolation is a mood buster [252]. Indeed, social activities trigger the release of dopamine [245] (which we will discuss in more detail in subsection 3.3.3), which is the hormone of reward and provides us with instant pleasure that lasts for a short duration.

We are also beings who have evolved to be concerned about the opinions of other members of our tribe [264]. Therefore, we are beings that require social validation [330][264], even from a computer, as we will elaborate in the following subsection 3.3.2 [161]. Additionally, we have a tendency to enjoy talking about ourselves [245]. Social networks provide us with all of this, and they have been designed to make us dependent, emotionally vulnerable, and easily manipulable, as we will explore further later on.

3.3.2. How social networks create the illusion of being alive

Have you ever heard of persuasive technologies? These technologies have a specific mission: to modify people's behavior [160] and encourage them to spend more time on their platforms, as is the case with social networks [264]. Companies such as Facebook, Google, and others have dedicated departments solely focused on persuasive technologies, with the objective of making users increasingly addicted to their products and attracting more individuals to their platforms [264]. These individuals are commonly known as *growth hackers*⁵. Chamath Palihapitiya was part of one of these departments at Facebook. These growth hackers are trained to understand how the human brain works, its flaws, and how to **gradually** and **imperceptibly** control it. Tristan Harris explains that there are courses at Stanford where they learn how to translate human persuasion techniques into technology. Sandy Parakilas adds that many people working in Silicon Valley (Facebook, Uber, LinkedIn, etc.) have taken these courses. Persuasive techniques are largely based on the theories of Dr. Brian Jeffrey Fogg, a social scientist and professor at Stanford University [264].

⁵Growth hacking is a set of techniques, or rather a mindset, a methodology, aimed at rapid growth, by any means, of a service or product [334]. A hacker is someone who uses unconventional means to access a system. The growth-hacker employs unconventional means to develop a user base. They can work on all stages of the customer acquisition process (acquisition - conversion - retention) to achieve their growth objectives [291]

According to Dr. B.J. Fogg, as explained in one of his books [161] published in 2002, the fact that we have social reactions to computer products, perceiving them as if they were alive is normal and automatic. It is written in our genes to react to signals from our environment, particularly to things that appear to be alive (as mentioned in subsection 3.3.1). These reactions are instinctive and therefore uncontrollable; they do not stem from rationality. When we perceive a social presence, even if it is fictional as in the case of a computer object, it can evoke emotions such as sadness, anger, empathy, and more. In his book [161], Dr. Fogg presents the five types of social cues that a computer can use to influence us:

1. Principle of Attractiveness: A computing technology that is visually attractive to target users is likely to be more persuasive as well.
2. Principle of Similarity: People are more readily persuaded by computing technology products that are similar to themselves in some way.
3. Principle of Praise: By offering praise, via words, images, symbols, or sounds, computing technology can lead users to be more open to persuasion.
4. Principle of Reciprocity: People will feel the need to reciprocate when computing technology has done a favor for them.
5. Principle of Authority: Computing technology that assumes roles of authority will have enhanced powers of persuasion.

For the principles of similarity and reciprocity, Mr. B.J. Fogg conducted experiments, the details of which are elaborated in Appendix B.4. We recommend referring directly to their report for a comprehensive understanding.

Here is our analysis of these experiments. The first two studies on similarity lead us to realize (as Mr. B.J. Fogg emphasizes) that the more users identify with the product, the more likely they are to be persuaded to change their attitude or behavior in the direction suggested by the product. It also suggests that people are more persuasive when interacting with a computer that appears similar to them, whether in terms of personality or affiliation. If we combine the principle of similarity with other persuasion techniques derived from psychology, computers would be capable of eliciting emotions such as joy, fear, anger, etc. They would also be able to negotiate and support ideas.

When focusing on social networks, everything is designed to align with our ideas and emotions

by creating products that resemble us (through the use of filter bubbles). As explained earlier, on Instagram and TikTok, the reels⁶ that are proposed to us are the ones that best match our emotions and interests (which they have understood), in order to offer us content that is similar to who we are. This makes the service feel more alive and enables them to manipulate us.

We observe that these studies were conducted in 2002 and yielded remarkably impressive results. We suspect that with the proliferation of social networks, the vast amount of data they have access to, and the undisclosed experiments they have conducted, they have developed highly sophisticated persuasion techniques.

3.3.3. Every action is designed to give you a shot of dopamine

Have you ever had the feeling of wanting something that you thought would bring you pleasure, but when you obtained it, the effects of pleasure vanished after just a few seconds? This pattern can be observed in several cases, some of which you will find familiar: chocolate, sweetened drinks, compulsive purchases, gambling, drugs, sports, sex, tobacco, and many others [13][77][5].

The link between the elements given in the list above is that they all follow the same pattern: we take a small dose and it gives us pleasure for a short time, i.e., a reward [5]. Then, we feel like we want more, so we take it again. The feeling of wanting more always becomes more and more intense until we no longer know how to satisfy this craving, resulting in a sense of emptiness and unease [Tristan Harris & Chamath Palihapitiya in the Netflix documentary [264]]. But why does this pattern occur? It is closely related to dopamine, often referred to as the pleasure hormone. Dopamine is a neurotransmitter that plays a crucial role in various functions such as motor control (including kidney and heart function), attention, sleep, memory, cognition, motivation, pleasure, addiction, love, and sexual pleasure [13][77]. When our brain associates a particular experience with pleasure, it releases dopamine, which in turn influences our behavior [5].

Dopamine is not the only hormone associated with pleasure. There is also serotonin [13][163].

⁶Reels refer to brief video clips, usually lasting less than one minute. Unlike the traditional scrolling experience found on platforms such as Facebook, where users navigate through different types of content, such as text, images, and articles, scrolling through reels entails an endless stream of concise videos. This format provides users with a continuous flow of captivating video content [333][6][97].

In Appendix C.3, we explain the difference between these two hormones, and we recommend reading it to gain a better understanding of the impact that social networks, which have become addictive due to the influence of dopamine, have on society. It is important to note that dopamine provides instant pleasure and satisfaction (excitement in the short term), while serotonin is responsible for long-term happiness (contentment in the long term) [13][163]. Our choices in life are often influenced by these hormones, and we tend to seek out dopamine, which can become addictive. An excessive release of dopamine or an imbalance between dopamine and serotonin can have various consequences, including an increase in aggressiveness [13][77].

Therefore, it is not without reason that we are witnessing a rise in global social media consumption: social networks have become a drug, reducing our ability to concentrate and think critically [287], leading to increased aggressiveness and impulsivity. We react instantaneously to what we see without giving it a second thought.

Notifications, like buttons, infinite feeds, shares, comments, the use of a red color for notification icons, the pull-to-refresh mechanism (resembling a slot machine where the outcome is unknown), and so on. Every aspect is meticulously designed to provide you with a small dose of dopamine **each time** [287][264]. On Instagram, when you post a photo, people can react. The particularity is that when you open Instagram, it takes 0.5 to 1 second before it displays the number of likes and comments you have received on your post. This delay is not accidental. It is intentionally designed to create tension before revealing the number of reactions you have received. When it finally appears, it triggers a release of dopamine in your body. Naturally, you will check multiple times a day to see if people have reacted. This also aligns with what was mentioned in subsection 3.3.1 regarding the search for social validation. Tim Kendall⁷, Justin Rosenstein⁸, Sean Parker⁹, Chamath Palihapitiya, Roger McNamee¹⁰, Sandy Parak-

⁷Tim Kendall was the director of monetization of Facebook (2006-2010). Then he became the head of product (2012-2015) and after president of Pinterest (2015-2018). His LinkedIn: <https://www.linkedin.com/in/tim-kendall-6572/>.

⁸Justin Rosenstein was the product manager at Facebook (2004-2007), then the engineering manager at Facebook (2007-2008). He co-created the like button at Facebook [43]. His LinkedIn: <https://www.linkedin.com/in/justinrosenstein/> and website : <https://justinrosenstein.com/>.

⁹Sean Parker was the founding chairman of Facebook in 2004. Now he is a venture capital investor and philanthropist [14][167][91]. His LinkedIn: <https://www.linkedin.com/in/parkersean/>.

¹⁰Roger McNamee was one of the first advisors of Mark Zuckerberg and one of the first investors in Facebook. [172][50]

ilas¹¹, Loren Brichter¹², and many others concur on the significant issue of social networks addiction. Everything is intentionally designed to make you addicted [287][264]. Regarding social validation, a study conducted in 2013 by Dar Meshi, a postdoctoral researcher at the Freie Universität in Berlin, on 31 participants revealed that receiving positive feedback about oneself triggers a stronger activation of the nucleus accumbens, a brain region responsible for reward and dopamine secretion. This region is also activated by drugs, money, sex, and food, making it associated with addictive behaviors. Social networks capitalize on this knowledge [254][47][245].

In the documentary *The Social Dilemma* [264], Tim Kendall shares his personal experience as the former president of Pinterest. He reveals that while he was working on making Pinterest addictive during the day, he himself became addicted to it. In the evenings, when he returned home to his wife and children, he struggled to put his phone away and found himself scrolling through Pinterest or checking his emails.

This dose of dopamine not only aims to make us addicted but also to **manipulate us**. The fact that we receive a reward (a dose of dopamine) for a specific action (scrolling, refreshing the page, checking our number of likes, etc.) conditions us to perform that action until it becomes automatic [287]. It's similar to a dog that salivates when you touch the bag of kibble. From birth, it was encoded in his genes that he should salivate when he sees food. Over time, his brain learned that the sound of the bag indicates that he will receive food. He has been conditioned to salivate at the sound rather than directly at the sight of food. This phenomenon is known as **intermittent positive reinforcement**. [264]. Intermittent positive reinforcement is a procedure where the likelihood of a behavior occurring increases when a stimulus is added in response to that behavior [193][8][71]. The act of scrolling, refreshing the page, and similar actions provide us with rewards by triggering the release of dopamine, as evidenced by studies [314][264]. This constant release of dopamine creates a desire for more, leading us to instinctively repeat these actions [264]. This concept is further explored in Appendix B.11.9. Consequently, over time, we become unaware of our actions, performing them automatically. In addition to fostering user dependence, technology contributes to a phenomenon known as continuous partial attention, which reduces the ability to focus and possibly lowering IQ [287]. Indeed, a study conducted in 2017 [347] demonstrated

¹¹Sandy Parakilas was the platform operations manager at Facebook (2011-2012). His Linkedin: <https://www.linkedin.com/in/sandyparakilas/>.

¹²Loren Brichter is a developer how created the pull-to-refresh mechanism that was first used by Twitter [287].

through two experiments the phenomenon of brain drain induced by smartphones: the mere presence of consumers' smartphones can have a detrimental effect on two cognitive capacities, namely available working memory capacity and functional fluid intelligence¹³. The presence of notifications, the desire to know what's happening on our favorite social networks, and other factors, generate the urge to check our smartphones.

In addition to impacting our memory, we also observe a decline in our attention span. A study conducted by Microsoft on 2,000 participants revealed that in 2000, the average attention span was 12 seconds. However, in 2015, the same study indicated a decrease to approximately 8 seconds (even shorter than the attention span of a goldfish which is 9 seconds) [188][234][351][245]. During the same period, between 2000 and 2015, the emergence of multitasking and social networking took place [28]. But what exactly is attention span, and why has it decreased? Attention span refers to the length of time an individual can fully concentrate on a specific task without being distracted [341]. This decline can be attributed, in part, to our current consumption patterns. To come back to the influence that dopamine has on our brain, our brain is strongly influenced by our emotions triggered by various events. MRI scans reveal that there are two competing areas of our brain that influence our behavior and decision-making: the pursuit of short-term rewards and the pursuit of long-term goals. Can you see where this is leading? In general, we tend to choose instant gratification, which is associated with the release of dopamine [351]. So, what does this have to do with social networks (and technology in general)? They have simply conditioned us to be constantly bombarded with information, notifications, and so on, and to never have to wait for this information: every second, we receive our dose of dopamine.

To further highlight the way we have been conditioned to consume information quickly, to lose patience, and to crave multiple sources of information simultaneously, you will find in Appendix B.9 some figures that reflect our Internet consumption habits. The key takeaway from this appendix is that people expect things to happen quickly; we have lost our ability to be patient.

¹³General intelligence, which encompasses fluid intelligence and crystallized intelligence, plays a role in learning and memory. Fluid intelligence involves comprehension, reasoning, and problem-solving, while crystallized intelligence involves the retrieval of stored knowledge and past experiences [90]

4. When the system is beneficial to third parties

4.1. Advertising, the main source of revenue

Now that they have succeeded in getting us addicted to their platforms, incorporating advertising is crucial to complete their business model. Advertising serves as their primary revenue stream: the more people remain on their platform, the more advertising they can display, resulting in increased revenue and data collection to enhance their system. So, it becomes a cycle. Advertising represents 90% of Meta’s total revenue [339], amounting to \$115 billion in 2021 [304] (turnover: \$118 billion in 2021) [320]. Similarly, it constitutes 80% of Alphabet’s revenue [339][338], totaling \$209 billion in 2021 [335] (turnover: \$257 billion in 2021) [36][166]. You can find graphs illustrating these figures in Appendix E.2.1.

To achieve this, these companies offer an extremely sophisticated advertising targeting system based on the vast amount of data they have collected on their users. There is nothing inherently wrong with this system, as it is normal for companies like Facebook to provide advertisers with an advertising targeting system as part of their business model. However, it is entirely logical that such a system (including the platforms themselves and the advertising system) can be exploited by malicious individuals seeking to promote an ideology, divide society, or manipulate public opinion. Indeed, when considering the system as a whole, we observe a network of individuals often grouped into ideological clusters, facilitated by the creation of echo chambers and filter bubbles (as discussed in Chapter 5). This contributes to a mental and psychological weakening of people (as explained earlier), etc. Furthermore, advertising targeting tools allow them to have a complete control over the individuals they intend to target and the number of people they aspire to reach. (refer to Appendix E.2.2 for visual examples

of the available settings for advertising on Facebook). While it is important to recognize the legitimate uses of advertising targeting systems, it is also necessary to become aware of the potential risks associated with such tools and their impact on society as a whole.

We will now examine a series of facts that have occurred on Facebook (and other platforms) and demonstrate the dangers of these platforms when third parties exploit them. We will discuss significant cases such as Cambridge Analytica, but we will also present lesser-known facts.

4.2. Cambridge Analytica

The entire story surrounding Cambridge Analytica that we are about to explain is based on numerous testimonies, investigations and interviews. We have noticed that there are divergences in the different stories that can be found. Therefore, we have taken into account multiple sources to reconstruct the facts as accurately as possible. As we will see, we realize that many events occurred, but providing a clear explanation of what happened can be challenging because these private companies (specialized in profiling and psychological manipulation) are highly cautious about keeping their internal operations secret. Additionally, they often own multiple companies, making it sometimes difficult to establish connections between them. In Appendix B.5.4, you will find a graph illustrating the structure of the SCL Group (the company that created Cambridge Analytica) to give you an idea of its complexity.

The story behind the creation of Cambridge Analytica is explain in Appendix B.5.1, where we also provide information on the company's investors and clients. What we will remember from this story is that Cambridge Analytica is primarily composed of individuals who seek to manipulate and exert power over the population. Their clients, some of whom have become members themselves, hold rather extreme and radical ideas. It has always been the case that parties with extreme ideas seek to gain power, control people, and adopt an aggressive discourse. With Cambridge Analytica (and the other stories we will share with you), we have come to realize that social networks are primarily beneficial and utilized by extremist parties who invest significant amounts of money in gathering data of people and utilizing the tools offered by platforms to propagate propaganda and manipulate people.

Concretely, what happened in 2014 and, by implication, in 2016? Let's go back to 2014

when, at the University of Cambridge, a team of researchers from the Psychometric Center has developed a technique to understand a person’s psychological profile solely based on their activity on Facebook, particularly based on what they like. Cambridge Analytica reached out to them for collaboration, but the research team declined, except for the psychology professor: Dr. Aleksandr Kogan [233][99].

In this collaboration, Mr. Kogan developed a new application called ThisIsYourDigitalLife in June 2014. This application, in exchange for monetary compensation, requested users to complete psychological tests and to grant the application access to their Facebook data, including their political beliefs, interests, and so on.[319][99]. At that time, granting a third party access to our Facebook data also allowed them to access the data of our Facebook friends (this functionality has since been deactivated) [63]. For instance, if an individual had 200 friends on Facebook and completed the form, the data of those 200 friends would also be retrieved. The data collected by Mr. Kogan was subsequently sold to Cambridge Analytica. In the end, more than 270,000 people downloaded this application, resulting in the retrieval of data from over 50-80 million profiles [233][100][319][124] (a detailed diagram illustrating the functioning of the application, data collection, and key players involved can be found in Appendix B.5.5).

In this case, there was no system flaw or hacking involved. They just advantage of the Facebook platform for their own interests. As an example, The New York Times published an email exchange between Mr. Kogan and Mr. Wylie¹ on May 9, 2014, where Mr.Kogan described the personal traits he could infer from the collected data (see Figure 4.1) [318]. Small aside: If they were able to deduce such a wide range of traits about a person from the collected data, it underscores the information that Facebook is capable of inferring about its users.

¹Christopher Wylie is a researcher and data scientist who previously held the position of Director of Research at Cambridge Analytica and SCL Group. In 2018, he collaborated with The Guardian and The New York Times as a whistleblower, revealing the exploitation of social media data by Cambridge Analytica. [37]

Figure 4.1.: An email from Dr. Kogan to Mr. Wylie describing traits that could be predicted. [318]



The revelations regarding the Cambridge Analytica scandal became public on May 17, 2018, when Christopher Wylie collaborated with The New York Times and The Guardian to expose the incident [233][100][318][37]. However, it is important to note that Facebook became aware of the data breach before it was made public, as they learned about it through the press in late 2015. Indeed, in December 2015, The Guardian published an article [181] revealing that Ted Cruz's campaign was utilizing a company that had harvested data from millions of Facebook users without their knowledge. It was discovered that Ted Cruz's campaign had made a transfer of at least \$750,000 to Cambridge Analytica (for more information, refer to Appendix B.5.5).

Following the revelation, Facebook initiated an investigation and sent letters to Cambridge Analytica and Mr. Kogan, requesting the deletion of the data [318][100]. Surprisingly, Facebook did not verify whether Cambridge Analytica had indeed deleted the data. They relied on trust without conducting thorough checks. It was only after the events became public that Facebook started taking it more seriously [100][318]. This highlights two important aspects: firstly, Facebook did not inform its users about the data breach, and secondly, if there is nothing at stake for them, they don't care.

4.2.1. Cambridge Analytica's role in Trump's 2016 electoral victory

Before that, Cambridge Analytica had been working with Mr. Cruz's team. However, when Mr. Trump won the Republican nomination, the Mercer family, who had become one of the financiers of Mr. Trump's campaign [268], withdrew their support from Mr. Cruz and shifted their support to Mr. Trump [290]. They insisted that Mr. Trump put Mr. Bannon as the head of his presidential campaign and utilize the services of Cambridge Analytica, in which the Mercer family had made significant investments [198][331]. Mr. Bannon was a strategic person because he was the vice-president of Cambridge Analytica and the executive chairman of the conservative newspaper Breitbart News since 2012 [208][294]. This gave Donald Trump the advantage of having both Cambridge Analytica and a media outlet with significant readership. Indeed, Breitbart News reached its peak in July 2016 with 31 million unique visitors to its website (approximately 10% of the U.S. population at the time) and 192 million page views [266]. The influence and impact of Breitbart News sparked numerous internal discussions within Facebook [230].

Cambridge Analytica had utilized the collected data to target specific audiences for advertising purposes, conduct fundraising calls, predict voter turnout, purchase \$5 million worth of TV ads, and determine where Mr. Trump should go to get the most support [318]. In addition to successfully targeting individuals on Facebook, they also used Facebook's *Lookalike* Audiences feature. This feature enabled them to request Facebook to target individuals who shared similar behavioral and demographic characteristics with the previously identified targets of an advertisement [269][84]. Examples of advertisements created by Cambridge Analytica can be found in Appendix B.5.7.

4.2.2. Cambridge Analytica's role in influencing the Brexit vote

For this subsection, we will provide you with the key points. A comprehensive analysis and the complete story can be found in Appendix B.5.2.

The two pro-Brexit parties, Leave.EU and Vote Leave (which included Boris Johnson), engaged companies to influence the public's vote in their favor. Leave.EU paid Cambridge Analytica [65][197][306], while Vote Leave paid the services of Aggregate IQ [197][306]. In the end, it was Vote Leave that obtained the official designation. But, it later came to light that

Aggregate IQ and the SCL Group (the parent company of Cambridge Analytica) are closely connected [197][306]. Thus, although Leave.EU and Vote Leave initially appeared separate, they ultimately relied on the services of the same group: the SCL Group [101].

4.3. Russia

The SCL Group has also been linked to Russia, specifically with the Lukoil group. However, there is no evidence that Cambridge Analytica directly worked for Lukoil. Nonetheless, certain documents indicate that in 2014, Lukoil received comprehensive briefings on Facebook, microtargeting, data analysis, and election disruption [98]. Additional information can be found in Appendix B.5.3.

4.3.1. The IRA: an enterprise manipulating crowds on social networks

Considering the complexity and extensive nature of this subject, we will attempt to provide a concise overview. However, if you wish to delve deeper, we encourage you to consult the sources referenced, particularly [56][263][79][150]. These sources offer comprehensive information about the internal operations of the IRA and provide detailed explanations of how certain troll campaigns were orchestrated. Additionally, they feature testimonies from individuals who have worked for these companies, as documented by The New York Times and The Washington Post.

The IRA is a company that was founded in 2013 by Mr. Evgeny Prigozhin.² [79][56][256][328].

This company, based in Saint Petersburg, is primarily focused on influencing public opinions, promoting Russia, and creating disruptions in other countries. To achieve these objectives, the company employs various techniques. They create fake YouTube videos [56] and produces articles [263], replicate websites identical to legitimate news outlets, create fake Facebook pages and Twitter accounts, fabricate interactions between accounts to simulate genuine conversations (for example, one account will comment on a post while two other accounts will

²He is a Russian from Saint Petersburg. He is a businessman with close ties to Vladimir Putin and is the founder of the Wagner mercenary group, which operates to protect Russian interests outside the country [343][316][55][23].

contradict the comment by providing numerous links to articles they have created) [79], they like, share, and comment on their own posts to create the illusion of support [79][263][56], and employ proxies to avoid being easily detected and banned by platforms like Meta, Twitter, and Google [56]. Their tactics demonstrate a thorough understanding of the subject.

This company has notably created trolls in Colombia by making people believe that a chemical plant had exploded (Twitter was flooded with tweets and fake YouTube videos were circulating) [56]. They also made people believe that there was an Ebola epidemic in the USA [263][56] or that an unarmed black woman had been killed by a police officer [56]. They also operate in Africa (by posting about black history, civil rights, and LGBTQ issues on Facebook and Twitter) [150], against Ukraine [56], in Russia to praise Vladimir Putin and Russia [79][263], to make people believe that Syria’s withdrawal of chemical weapons was thanks to Russia [263], etc. They also had a significant impact on the 2016 U.S. presidential elections, which we will discuss further in the next subsection 4.3.2.

Here we are only listing what they have done, but we must realize that they manage to have a real impact in the world thanks to their sophisticated techniques and their deep understanding of how social networks work. They create panic, disorder, fear, and frustration. They fuel polarization and generate anger among people. Through their various means of infiltration on social networks, as explained earlier, combined with the tools offered by these platforms, they are able to sow doubt and monopolize the platform with their ideas to the point of stifling genuine information and simply creating panic. As Mr. Roger McNamee states, they have understood how Facebook works: They find 100 or 1,000 people who are angry and afraid and then use Facebook’s tools to advertise to get people into groups. That’s exactly how Facebook was designed to be used [269]. This situation is highly concerning.

4.3.2. How the IRA influenced the 2016 presidential election in favor of Donald Trump

Russia, and more specifically the IRA, have orchestrated disinformation campaigns and have spread fake news during the 2016 U.S. presidential elections. This case, known as Russiagate, was led by Mr. Robert Mueller³. In a 37-page report [297], he provides a clear explanation

³He served as the Director of the Federal Bureau of Investigation (FBI) from 2001 to 2013. In 2017, he conducted an investigation into potential Russian interference in the 2016 U.S. presidential election

that Russia’s operation was aimed at supporting Donald Trump’s presidential campaign and denigrating Ms. Hillary Clinton [125][214][297][116].

Indeed, a federal grand jury in Washington DC indicted 13 Russians and 3 entities, including the IRA, as mentioned in the report [214][297][293]. However, there is no evidence suggesting direct contact between Mr. Donald Trump and his team with the IRA and Russia to carry out this manipulation [353][125][214]. Nonetheless, it is worth noting that Mr. Donald Trump has received significant funding from Russian investors [298], and Kremlin-linked actors have provided substantial financial support to Mr. Trump [131]. These circumstances remain subject to doubt.

Specifically, what have they done? The IRA has employed various tactics across different platforms to achieve their objectives. They have created multiple groups on platforms such as Facebook, Instagram, YouTube (Google), Twitter, Tumblr, and Reddit [312][302][185]. On each platform, their primary aim has been to criticize Ms. Hillary Clinton and support Mr. Donald Trump [125][302][132]. They have specifically targeted individuals affiliated with the Democratic Party, attempting to dissuade them from voting. They additionally focused on ethnic minorities (including Latinos, Muslims, and Christians), the LGBT community, and notably, the African-American community, urging them to boycott the elections [143][155][312]. To reach these target groups, the IRA has created social media accounts and websites, such as blackmattersus.com, blacktivist.info, blacktolive.org, and blacksoul.us [312].

Between June 2015 and August 2017, nearly 126 million [259][143][281][139][302] Americans have been exposed to IRA content on Facebook, and 20 million on Instagram [192][337], totaling 146 million individuals. The IRA’s strategy on Facebook was relatively straightforward: they have utilized advertising tools to attract people to their groups, where the propaganda activities took place. The IRA’s advertisements (approximately 3,000 to 3,500 ads associated with 470 fake accounts [147][281][237][139][313]) accounted for around \$100,000 [237][259][139][302][67][313][269]. These advertisements reached 10 million U.S. voters between 2015 and 2016, with 44% of total ad impressions occurring before the elections [147][237]. Their primary objective was to encourage American voters to click Like or follow IRA-created Facebook profiles and pages, as well as target individuals who visited their pages. Their strategy was successful as they have managed to directly reach 30 million Facebook users with their 80,000 posts [281][139][259][302][67][143][192][337]. Subsequently, these posts have

[277][122][49]

been propagated through sharing, likes, and comments, reaching 126 million users. They have also utilized non-political but commercial advertising tactics to encourage people interested in specific products to provide their personal data for purchases, allowing them to target these individuals with advertisements [235] (An example of this type of post is shown in Appendix B.6.1.2). Additionally, they have offered coupons for products in exchange for sharing posts on Facebook [132] (An example of this type of post is shown in Appendix B.6.1.8). Among these groups, the most followed ones are: United Muslims of America (300,000 followers), Don't Shoot Us (250,000 followers), Being Patriotic (200,000 followers), and Secured Borders (130,000 followers) [302][297][185]. In total, this amounts to at least 1.8 million distinct individuals who have followed one of their Facebook pages [192]. In Appendix B.6.2, you can find a figure that displays the top 20 IRA Facebook pages ranked by the total number of likes received. On Instagram, the IRA monitored 170 accounts [192][302], which collectively posted 120,000 posts [192][302][132] and received 185 million likes [235][284][132] (it should be noted that some people also view the content without liking it, so the audience is even larger) between 2015 and 2017. Although Facebook never disclosed the number of followers these pages had, a dozen of them attracted over 100,000 followers, with approximately 50 pages exceeding 10,000 followers [284]. For instance, @blackstagram_ has more than 300,000 followers [312][284][132], and @american.veterans has over 200,000 followers [235][132]. Each group focused on a different topic and targeted a specific audience, including African-Americans, gun-rights supporters, and anti-immigration activists. Their objective was to create a sense of community within these groups (examples of their posts can be found in Appendix B.6.1) [235]. These groups also served the purpose of promoting Mr. Donald Trump and spreading fake news against Ms. Hillary Clinton. For instance, the account @feminism_tag had accused Ms. Clinton of insulting women who had accused her husband, Mr. Bill Clinton, of sexual misconduct. Another post, targeting African-American voters, featured a manipulated image of Ms. Clinton with braided hair and the caption when you need the black vote [235]. Approximately 7% of the Instagram posts targeted Ms. Clinton, while 11% mentioned Mr. Trump [235][132]. Facebook and Instagram together would have generated 220 million engagements [132].

On Twitter, with their 4,000 fake accounts [231][302][250] (201 of which were linked to other Facebook accounts of the IRA [143][280]), their strategy was clear and double: create fake accounts and then using a botnet of automated accounts to amplify them. The IRA tweeted

public. The IRA tweeted publicly, but also sent direct messages to people [302]. Twitter also discovered that the IRA had between 35,000 and 50,000 bots (automated accounts) [139][143][259][337] that allowed them to amplify their posts (like, retweet, ...) [231]. Their Twitter posts would have reached 32 million distinct users. Between September 1, 2016, and November 15, 2016, they published 131,000 tweets [139][259], and the approximately 1.4 million tweets received 288 million impressions [259][139][192]. In total, over 10 million tweets were posted [132], generating 73 million engagements⁴ [132]. The main accounts (approximately 123 according to Symantec) consisted mainly of fake news sources posing as regional news outlets or political organizations, with most of them having at least 10,000 followers. They also had other accounts (approximately 3,713 according to Symantec) with fewer than 10,000 followers, whose primary purpose was to retweet tweets from other accounts, research findings, and so on [231].

On YouTube, they would have uploaded over 1,100 videos [259][132]. Google had identified 18 channels associated with the IRA, which collectively have received approximately 309,000 views (according to Google) [259]. The two main IRA channels were *Don't Shoot* and *BlackToLive*, which focused on the Black Lives Matter movement and police brutality (Russia 2016). In addition to the videos, Google reported receiving a total of \$4,700 in advertising from these channels[281][259][139], and \$53,000 for political advertising from Russian accounts [139][259].

The IRA demonstrated a well understanding of how social network algorithms work. They effectively have utilized targeting tools to reach specific individuals and created clusters of like-minded people to manipulate them more easily. As we will discuss in chapter 5, social networks tend to facilitate the formation of echo chambers and use filter bubbles. Additionally, they prioritize viral content that generates significant engagement. The IRA has used this knowledge by deploying bots to amplify their content and create the illusion of a supportive community around their information.

In Appendix B.6.1, you will find several examples of ads created by the IRA. If you are interested in learning more, we recommend you to consult the following insightful reports: [132][185].

⁴Engagement refers to the number of people who viewed the content, while impressions indicate the number of times the content was displayed [64]

4.4. Sophie Zhang : a whistleblower's account about fake engagement in politics

Ms. Sophie Zhang worked at Facebook for 2.5 years and was fired in mid-August 2020. She denounced several facts about fake accounts that have been created to support political leaders worldwide, as well as internal issues within the company that went unaddressed. In this section, we will examine the various facts she highlighted. This section is mainly based on these two sources [219] (which contains a testimony of Ms. Sophie Zhang) and [221].

4.4.1. Examining Facebook's inactive response to election manipulation in Honduras

Honduras, a country located in Central America, held its presidential election in 2017 for the period 2018-2022. In the end, Mr. Juan Orlando Hernández was re-elected, having previously won the election in 2013 for the period 2014-2018 [15][162]. After 6 months of working at Facebook, Ms. Zhang discovered that Mr. Hernandez's page has received a significant number of reactions from fake accounts. In fact, out of the 59,100 distinct users who reacted to his posts between June and July 2018, 78% was fake people. These accounts and pages were created in 2016 and 2017, preceding Hernandez's disputed re-election. What was very clever of Mr. Hernández was that these fake people were fake pages and not fake accounts⁵. Indeed, Facebook's policy prohibits the creation of fake accounts and aims to ensure the authenticity of individual profiles. However, these rules did not extend to Facebook pages, allowing individuals to create multiple pages using a single account⁶.

This is precisely what Mr. Hernandez did. The account managing his official page also owned numerous fictitious pages. This enabled them to publish content on Hernandez's page and utilize the hundreds of other pages to react (like, comment, share) to posts, creating the illusion of popularity (the digital equivalent of bussing in a fake crowd for a speech). This technique is similar to the tactics used by the IRA during the 2016 U.S. elections, where

⁵An individual can create multiple pages using fake names and profile pictures, allowing them to perform the same actions as they would with a real profile. For instance, if someone has Page A and Page B, they can like, comment, and share the same post using both pages.

⁶From a Facebook report [349] : Facebooks current approach to addressing account integrity focuses on the authenticity of the accounts in question and their behaviors, not the content of the material created

they employed fake accounts to manipulate the population and influence political discussions.

4.4.1.1. When Facebook thinks it's not important

After the events of the 2016 U.S. elections, Facebook implemented a policy known as Coordinated Inauthentic Behavior (CIB)⁷, which is a standard used to ban accounts that do not comply to it, such as those involved in fake engagements.

However, Facebook initially chose not to classify the activity in Honduras as CIB for two main reasons. Firstly, as mentioned earlier, the organization was utilizing fake pages instead of fake accounts, which meant that it did not fall under the same regulations. Secondly, given that between 3 and 5 million individuals (depending of the sources [323][221]) in Honduras were using Facebook at the time, it appears that this issue was not a priority for the company. Ms. Zhang explained that nobody in the office seemed interested in investigating further, and she received messages such as: *I don't think Honduras is big on people's minds here* from a manager in the civic integrity team.

Ms. Zhang persevered, and on July 25, 2019, Facebook banned 181 accounts and 1,488 pages associated with this activity in Honduras. The individuals responsible for these pages had spent over \$23,000 on Facebook ads, indicating their utilization of Facebook's advertising tools. However, it is worth noting that Facebook's policy dictates that if fake Facebook pages are banned, the authentic account associated with them will not be banned. Consequently, the account holder can simply directly create new fake pages.

In the subsequent days, she continued her investigation and uncovered fake pages supporting political leaders in various countries, including Albania, Azerbaijan, Mexico, Argentina, Italy, Philippines, Afghanistan, South Korea, Bolivia, Ecuador, Iraq, Tunisia, Turkey, Taiwan, Paraguay, El Salvador, India, the Dominican Republic, Indonesia, Ukraine, Poland and Mongolia.

However, they did not invest significant time and resources in all of these countries, as we will see in the subsequent subsection 4.4.2.

⁷Coordinated inauthentic behavior (CIB) is a manipulative communication tactic that uses a mix of authentic, fake, and duplicated social media accounts to operate as an adversarial network (AN) across multiple social media platforms [261][262].

4.4.2. Facebook's prioritization of Coordinated Inauthentic Behavior (CIB)

Ms. Zhang began to voice her complaints about Facebook's lack of action regarding certain instances of Coordinated Inauthentic Behavior (CIB) and had multiple conversations with Mr. Guy Rosen, the Chief Information Security Officer since 2013. During these discussions, Mr. Rosen explained that Facebook faced challenges in addressing all issues and had to prioritize the most urgent matters. He stated that combating false likes was not a top priority for them. According to him, their main focus was on addressing issues in the United States, Western Europe, and countering foreign adversaries such as Russia, Iran, and others. He also emphasized that Facebook did not have unlimited resources to tackle all of these problems. This is surprising considering the net revenues they generate.

One of the only factors that could lead Facebook to remove these prioritization measures is if the press or the public gets involved, as Facebook cannot ignore the issue in that case. Therefore, it is interesting to note that even if they detect Coordinated Inauthentic Behavior, if it goes unnoticed and is not considered a priority, especially in poorer countries, Facebook will take no action. For instance, Ms. Zhang recounts that in September 2018, she informed the political team in Mexico City about Alejandro Murat Hinojosa, the governor of Oaxaca, that is was receiving a significant number of fake engagements on Facebook. However, it was not until January 7, 2019, that Facebook took action after receiving an article from the Atlantic Council's DFRLab that exposed suspicious likes from South Asian Facebook accounts on Murat's Pages. Ms. Zhang was assigned to this case and discovered 18 networks engaged in similar activities with local politicians, which were not prioritized by Facebook and therefore, remained active until August 2020. Later on, 5,000 accounts were finally banned.

4.4.3. Powerlessness of Facebook in Azerbaijan due to a lack of moderators

The case of Azerbaijan was the most complex and time-consuming for Ms. Zhang, where fake accounts and pages were employed to support an authoritarian regime with an egregious record on human rights

Azerbaijan employed a similar tactic as in Honduras, whereby they created fake pages using a

there were still 200 suspicious accounts promoting a politician in Bolivia, 100 in Ecuador, 500 in Brazil, 700 in Ukraine, 1700 in Iraq, 4,000 in India, and over 10,000 in Mexico. Furthermore, she noted the presence of approximately 130 recidivist users who remained active to support the President Juan Orlando Hernández.

For your information, Meta regularly publishes the accounts and pages that they ban on their official website: [link](#). It is important to note that while they detect cases of coordinated inauthentic behavior, it does not necessarily mean that they will allocate resources to dismantle them.

4.5. Facebook's amplification of hate speech and disinformation in Myanmar

To provide some context, in Myanmar, the predominant religion is Buddhism, practiced by 90% of the population [85]. There is a minority Muslim group called the Rohingyas residing in the state of Arakan. In 1982, the Rohingyas were stripped of their Myanmar citizenship, claiming they were not present in the territory before the British invasion in 1823. Since then, they have faced severe discrimination, including the loss of voting rights and limited access to certain jobs [222][325].

Starting from 2016, the Burmese government launched a violent campaign targeting the Rohingyas, involving massacres, rape, and torture [58]. The Rohingyas reacted, making the situation even worse [325]. As a result, nearly 9,000 Rohingyas were killed, and approximately 700,000 to 900,000 Rohingyas fled to Bangladesh [222][325][58].

It is important to note that economic interests likely play a role in the government's actions, as the Arakan region is rich in natural resources such as gas, oil, timber, precious metals, and minerals [222].

It should be noted that the government certainly has economic interests, as Arakan is a region rich in gas, oil, timber, precious metals and minerals [222].

Now that the context is established, let's examine the role played by Facebook in this conflict. It is important to note that in developing countries, Facebook serves as the primary source of information and a key communication platform [224]. In Myanmar, many people use Facebook

as their only gateway to the internet to access news and information [151][224].

During the crisis, hate speech proliferated on the platform. Facebook played a significant role in amplifying this hate speech [25][199]. For instance, Raymond Serrato and digital analysts examined 15,000 Facebook posts from the Ma Ba Tha group, an anti-Rohingya group. They discovered that the initial posts were made in June 2016, and there was a significant increase in activity on August 24th and 25th, 2017, following the counteraction by the Rohingya. This particular group, which had 55,000 members, saw its interactions increase by 200% [241].

Other researchers made other discoveries: they started liking the Myanmar army's fan page (which was not considered to violate Facebook's rules). Afterwards, Facebook recommended them to follow pro-military groups that contained abusive contents. They didn't have to wait long before Facebook redirected them to these groups, because of the first five groups recommended by Facebook, three contained information that went against the rules of Facebook, as glorify and incite violence. [151]

There was also a lot of disinformation against the Rohingyas and the Muslim religion, such as the fact that they store weapons in their mosques [241]. Facebook has studied the subject and admitted that its platform has been utilized to incite offline violence. They have admitted that the functionality of Facebook, including features such as virality, recommendations, and engagement optimization, plays a significant role in the proliferation of hate speech on the platform [224][25][272].

4.5.1. Facebook prioritizes its economic growth at the expense of its responsibilities

There are multiple factors contributing to this situation. Firstly, Facebook has consistently prioritized economic growth, leading to a neglect of human rights considerations. It was not until 2014, when the media started talking about the situation, that Facebook started to investigate the issue [272]. As mentioned earlier, as long as there is no public discourse on the matter, Facebook will not allocate resources to address the issues, even if they are aware of them. Brian Boland, a former vice president of Facebook responsible for partnerships with ISPs in Africa and Asia, expressed his opinion that Facebook prioritized investing in security measures primarily in developed countries such as the USA and Europe, while other countries

were merely used to increase user numbers. They did not invest in developing countries, despite having a larger user base compared to developed countries, where social networks are often the main source of information and communication [224].

Another point highlighted by The Wall Street Journal is that in 2020, a Facebook document revealed that employees spent over 3.2 million hours labeling information as true or false. However 87% of this information came from the U.S [224]. This demonstrates Facebook's limited interest in the information circulating on its platform outside of the USA...

A second reason is the persistent issue of moderation and language. Indeed, in these countries, the Facebook moderators are almost nonexistent or do not speak all languages of the country. Furthermore, Facebook heavily relies on algorithms to detect fraud, hate speech, and other content violating their rules. However, the effectiveness of their artificial intelligence is compromised by the diverse range of languages and the fact that it does not cover all linguistic variations [224][126]. Two internal teams at Facebook, estimated that the AI was deleting posts that generated between 2% and 5% of the views of hate speech and 0.6% of all content that violated Facebook's policies against violence and incitement [126] (that's not a lot). In the following section, we will see a similar scenario that happened in Ethiopia, where Facebook employees acknowledged that their information classification AI did not work because it is not adapted to all cultures [224]. For instance, during the conflict between Israel and Palestine, Facebook inadvertently removed Arabic-language regional news sources, activists, and posts containing the term AL Aqsa (referring to a mosque in Jerusalem). These deletions were the result of the AI's erroneous classification, as AL Aqsa is also associated with the Al Aqsa Martyrs Brigade, which is recognized as a terrorist organization by the United States [303][224]. It is worth noting that Facebook moderators handle between 50 and 70 languages (50 at the time of the events in Myanmar) [103][224], whereas there are approximately 7,000 languages spoken worldwide (although about twenty languages are spoken by 95% of the population) [326]. It is also important to consider that languages can have variations, as seen with Arabic, where different dialects exist. Facebook has acknowledged that its AI and moderators do not fully comprehend all the diverse Arabic dialects [246][224][102]..

In November 2022, Facebook laid off 11,000 employees, which accounted for 13% of its total workforce, due to poor economic results [244][310][61]. They want to fire again 10,000 employees by the end of May 2023 and close 5,000 job offers [271][336][258]. However, the specific types of employees who were laid off were never communicated. While it is likely that

engineers and programmers were among those laid off, it remains unclear whether moderators were also included. If moderators were indeed laid off, this would further diminish their ability to effectively manage the platform. Similar situations have been observed in other large companies as well (Twitter[119], Microsoft[216], ...).

4.6. Facebook, a platform for hate and violence in Ethiopia

To provide context, although tensions began before 2020, it was on November 3, 2020, that the conflict in Tigray began in the country. The Ethiopian government accused the Tigray People's Liberation Front (TPLF) forces of attacking the Ethiopian federal army based in Tigray and therefore decided to launch an offensive in that region. However, several regions provided military support to the federal government [21]. This war, which ended in November 2022 [86], resulted in the deaths of around 600,000 civilians, the displacement of 2.6 million people, and an increase in humanitarian needs in the country [26][190][21].

Facebook's role in this conflict was quite similar to its role in Myanmar: they did nothing to prevent the spread of hate. The Ethiopian armed forces used Facebook to incite violence against minority groups (the Tigrayans) [144][224], while the TPLF (Tigray People's Liberation Front) also utilized the platform to gather supporters. So there was a virtual war on Facebook. Facebook became aware that Ethiopia was a high-risk area on its platform [144]. However, they let posts that incited violence, spread misinformation, etc. [200].

In Ethiopia, Facebook has engaged third-party companies, such as PesaCheck and AFP, to do fact-checking [144]. These companies have teams dedicated to fact-checking in Ethiopia, although they are not based in the country [144][200]. The content that was determined to violate Facebook's rules by these two companies has been removed from the platform. However, another fact-checking company named HaqCheck (which had proposed a partnership to Facebook, but their offer was refused) had identified posts as violating Facebook's rules, but these posts remained on the platform [200]. The problem lies in the fact that the dissemination of fake news on the platform has resulted in the loss of multiple lives, such as the death of 81 individuals in the Oromia region [151][4]. There was also death of Professor Meareg Amare Abrha, who was targeted due to posts that identified him and incited hatred on Facebook.

Despite numerous reports made by individuals before the incident, the posts were only removed after the professor's death [109]. So, it's the same situation as in Myanmar where Facebook took no action, lacked the necessary tools to address the issue (with misclassifications of posts by AI systems [224]), and played a role in amplifying the conflict.

4.7. Facebook's ad targeting tool strongly utilized in Europe

Before starting this section, here is an interesting link where you can find studies on Facebook advertising in Europe: <https://adlens-be.medium.com/>. It should be noted that these studies rely on Facebook's database, which means that only the ads labeled as political advertising by Facebook are included. It is important to keep this in mind because when you pay for advertising on Facebook and declare it as a political advertisement, Facebook will target a different audience. If you don't declare it, Facebook will analyze your ad using an algorithm. However, a study conducted by researchers at KU Leuven and New York University has shown that Facebook's detection algorithm is incorrect in 83% of cases. It struggles to identify when an ad is a political advertisement and often categorizes non-political ads as political ones [51][239]. It appears, therefore, that bypassing this firewall to target an audience that is not supposed to receive this type of advertising is relatively straightforward ⁸.

4.7.1. Extremist political parties, the biggest spenders in social media advertising in Belgium

In Belgium, political parties have the right to advertise on social networks. However, there is a significant difference in the amount of money spent on advertising among the various political parties.

In 2020, the total amount spent on Facebook advertising in Belgium by all political parties was 4,858,980 euros. If we take a closer look, we notice that 78% of this amount comes exclusively from 3 parties: NV-A (1.48 million euros), Vlaams Belang (1.43 million euros), and PVDA

⁸Note: You can view the advertisements from various Facebook pages and get an estimation of the amount they spent on each post at the following link: <https://www.facebook.com/ads/library/>

+ PTB (0.88 million euros). In comparison, the PS spent 20,000 euros, and the MR spent 131,000 euros [81]. In 2022, a total of 4,999,583 euros was spent, of which 71% was contributed by the same 3 parties [82]. We can observe that extremist parties are the ones spending the highest amount of money on these targeting tools.

Another interesting aspect to note is that Vlaams Belang decided to launch its own online media platform in 2020: VBTV, where we can find headlines such as *Turkey and its negative influence on our country and Europe*. Reinout Van Zandyeke, an expert in political communication, explains: This initiative is aligned with American media outlets like 'Breitbart' or 'One America News'. A study reveals that supporters of Donald Trump and Joe Biden follow different media sources, indicating an increasing polarization within society. However, this can be potentially dangerous as individuals who rely on such media sources should be cautious not to mistake them for traditional news outlets. It is important to recognize that these platforms primarily provide opinions rather than objective and neutral information. Moreover, Vlaams Belang has also embarked on campaigns through WhatsApp, following a similar approach employed by Jair Bolsonaro during the presidential election campaign in Brazil [83].

4.7.2. Extremist Parties seem to invest heavily in Facebook advertising across Europe

For Europe, you will find our analysis in Appendix B.7. Unfortunately, we did not have the opportunity to carry out a complete analysis. The appendix primarily includes information on the parties and countries that have invested the most money in advertising, as well as an analysis of Hungary and Italy, where far-right parties have recently achieved success.

In Hungary, there is a network of political influencers called *Megafon* that circumvents the rules on political campaign advertising. They are closely associated with the Fidesz political party (Mr. Orban is a member), which they promote on Facebook [329][345][10]. In Italy, Georgia Meloni of the Fratelli d'Italia party emerged as the winner in 2022. It was one of the parties that spent a significant amount of money on Facebook advertising [251].

4.8. Non-exhaustive list of instances of negative impact by social networks

We will now provide you with a non-exhaustive list of instances where social networks have played a negative role. We will provide less detailed information as a comprehensive analysis would require a larger document. However, feel free to refer to the various sources if you want to know more.

- **India.** Facebook's recommendation algorithm incites violence, spreads misinformation, amplifies violent and anti-Muslim posts in India. Facebook is clearly aware of this. Facebook researchers have also discovered that the algorithm recommends groups with anti-Pakistan posts and suggests pictures with deceased individuals. In 2020, Facebook has 400 million users in India (its largest market) [108]. Once again, Facebook plays an amplifying role in the dissemination of false hate information and fails to manage it. See: [265] [317][105].
- **Drug cartels.** Mexican drug cartels have used Facebook to recruit, train, and pay hitmen. See: [224].
- **Human traffickers.** Employees (of Facebook) reported that human traffickers in the Middle East were using the platform to lure women into abusive employment situations where they were treated as slaves or forced into the sex trade. See: [224].
- **Bad posts's classifier for Hindi and Bengali.** Hindi is one of the most widely spoken languages in the world and one of the official languages of India. Along with Bengali, it is one of the two most spoken languages in India [128][44]. Facebook has acknowledged that they introduced content classifiers (detecting incitement to violence, etc.) for these languages in 2018 (for Hindi) and 2020 (for Bengali). Prior to that, there were no tools to detect misinformation, incitement to hatred, and so on. See: [105][265][344][273].
- **Censorship in Vietnam.** In 2020, Facebook was faced with a choice: the (pro-communist) government told them that their platform would be shut down in the country if they did not censor (and therefore over-punish) anti-government posts. This deal was worth 1 billion dollars for Facebook, so Mr. Mark Zuckerberg accepted. Funny for a platform promoting freedom of expression, except when there is money at stake.... See: [145][344].

- **Election in Brazil - 2018.** WhatsApp is the most widely used messaging application in this country. As a result, it is used to spread rumors, fake news, and disinformation. A study has shown that out of a database of 100,000 political photos circulating in 375 groups, among the 50 most widely propagated photos, 28 were misleading [117]. Moreover, the majority of fake news propagated on this platform favored Mr. Jair Bolsonaro. In fact, another study collected 11,957 viral messages shared across 296 groups, and among these messages, 42% of right-wing items were found to contain false information by fact-checkers [120]. There are also many other interesting events and information such as the fact that Youtube was also used for disinformation [69]. See: [117][120][69][78]. In 2022, the presidential elections took place, and this same pattern has resurfaced (Facebook and Google have not learned from the past). Indeed, their platforms have been reused to propagate fake news and misinformation. The human rights organization Global Witness (an NGO) wanted to test the difficulty of placing misleading advertisements on Facebook and YouTube. They voluntarily created a series of misleading ads. Among these ads, half of them were accepted by Facebook, and all of them were accepted by YouTube (the NGO removed these ads before the platforms broadcasted them) [74]. See: [121][74].
- **Human traffickers.** Social networks are heavily used for human trafficking (especially for the sale of domestic workers), and Facebook is aware of this and admits that they are not prepared to counter it on their platform. The issue has become so significant that Apple has threatened to remove the Facebook application from their App Store. Facebook swiftly removed the problematic content from its platform. Although Facebook managed to stay on the App Store, the problem persisted and even worsened as the trafficking also extended to Instagram. We have provided a brief explanation, but the issue is extremely serious. Please refer to the following sources for more information: [113][283][134][344] which provide detailed information.
- **XCheck.** Facebook has a program known as Cross Check or XCheck, which consists of a list of VIP individuals including celebrities, journalists, politicians, and others (with 5.8 million accounts on this list in 2020). This list allows certain individuals the ability to violate Facebook's rules without facing any consequences. For instance, one of the whitelisted accounts shared false and inflammatory claims, such as the assertion that vaccines are lethal, that Ms. Hillary Clinton had concealed pedophile rings, and

that then-President Donald Trump referred to all asylum-seeking refugees as animals. Another example involves the footballer Neymar, who was accused of rape in 2019. Subsequently, he posted explicit photos of the woman and shared WhatsApp conversations with her name on Facebook and Instagram. These photos were simply removed by Meta without deleting his account because he is in the XCheck list. An alarming statement from a confidential Facebook review stated : *We are not actually doing what we say we do publicly*. See: [203].

- **Criminals use platforms of Meta to buy and sell children for sex.** Meta claims that they are making every effort to address the issue but an two-year investigation have seen evidence that suggests it is failing to report or even detect the full extent of what is happening, and many of those victims they interviewed said they felt powerless to get the company to act. For more information, please see this article from The Guardian: [229].

We have observed that in various instances, Meta teams have acknowledged the negative impacts of their platform in areas such as mental health, human trafficking, and incitement to hatred. However, it is noteworthy that if no one raises the issue, they take no action to address it. And when the media address these issues, it does not necessarily mean that Meta will make an effort to resolve them. The damage is generally already done and has become so widespread that it is difficult to stop it. The most important thing to remember is that Facebook operates based on its business interests. If their platform generates interaction, even if it involves violence, it remains profitable for them. Sometimes they don't solve the problem because it's time and money-consuming, which affects their business [203]. Within the company, several employees have already attempted to propose new approaches to restrict features that disproportionately amplify incendiary and divisive posts. However, the company has refused these proposals, citing potential negative impacts on their growth. Indeed, Mr. Zuckerberg has frequently emphasized the advantage that virality brings to the economy of his company, which relies on their ability to encourage users to return to the platform by maximizing the viral spread of messages via the sharing button system [204]. John Hegeman, Vice President and Head of Monetization at Meta [42], states that *If we remove a small percentage of reshares from people's inventory, they decide to come back to Facebook less* [204].

5. Fake news

5.1. Difficulties of the study of fake news

The study of fake news on the Internet and social networks is quite vast, but all of them can be subject to criticism. This criticism does not stem from a lack of rigor, but rather from the two main challenges that we will explain. Although the conclusions of some articles and this thesis are strong, it is essential to maintain a critical mindset and not hesitate to question certain points, which are often addressed within the articles themselves.

One of the challenges is simply to define what is a fake news? It is crucial to establish a clear definition as it impacts the outcomes of various studies. Indeed, if a study's methodology involves labeling data as either true or false, the classification of data can vary depending on the definition used. Consequently, it can change the result of the studies. More information regarding the definition of fake news can be found in the Appendix C.1.

The second difficulty lies in obtaining qualitative data. As mentioned before, there is an asymmetry of information: social networks are the only ones who know what's going on on their platform. Because the data is private, it prevents researchers from conducting studies and observing what actually happens. Indeed, the results of studies depend on their databases. Most social network databases are private [68][170], so research is often based on a public database that does not objectively represent reality [301]. In fact, there is always a subset of data that favors a negative or positive result for the research question. Moreover, researchers are left in the dark about when and how social media providers change the sampling and filtering of their data streams. This is why Europe wants social networks to share their databases for research purposes [213][301].

5.1.1. Social networks: a platform conducive to fakes news

Social networks are complex systems. It is necessary to take into account several parameters and the implications of these parameters in order to have a comprehensive understanding of this system that facilitates the spread of fake news and makes them more visible compared to real objective information. The three main parameters are: the cost of disseminating information, the creation of clusters and echo chambers, and the type of information that circulates. We will discuss these parameters by attempting to establish the most insightful and accurate connections between them.

5.1.1.1. The cost of disseminating information has become free

After the First World War, journalistic standards of objectivity and balance were established by journalists who were confronted with the use of propaganda during that war. The arrival of new newspapers and technologies such as mass printing and radio did not prevent the maintenance of these standards [238][53]. However, with the advent of the Internet and, particularly, social networks, the cost of disseminating information has become free [238][68]. Anyone can now proclaim themselves a journalist, create a blog, a YouTube channel, a Facebook group/-page, and so on, and freely share content on any topic without having to substantiate the veracity of their claims. Anyone can post, share, comment, and like content that aligns with their own beliefs. However, most of us lack the qualifications of professional journalists and do not respect the standards of objectivity mentioned earlier [238]. Indeed, it's human nature to want to defend its values and community and sometimes going to extremes against objectivity (we'll come back to this in the following subsection 5.1.1.2).

5.1.1.2. The creation of clusters and echo chambers

Human beings need to have the feeling of belonging to a community. This feeling of being in a group with people who share the same values and ideas gives them a sense of strength, support, belonging, and the belief that their ideas are correct because they are validated by multiple individuals. Social networks allow people to gather in groups, pages, and other online communities. We also have YouTubers who have communities with their viewers, who sometimes even give themselves a name (e.g., PewDiePie with his *Bro Army* [38] community

and 111 million followers). Social networks allow us to join groups and communities, but above all, they guide us towards them. Indeed, based on the profile they create about us, they will suggest groups and content that align with our ideas and interests. This promotes the formation of clusters and echo chambers [127][177]. This means that, whatever ideas we have, we will always be guided towards communities that share the same values, and therefore we will be less likely to question them because we feel that our ideas are validated by others. Additionally, we can also add the concept of filter bubbles [177][92][68], which are created by social networks or more broadly by Google. We have already indirectly referred to this concept in this thesis. This concept was theorized by Eli Pariser, who say that we are intellectually isolated from others. This isolation is due to the recommendation algorithm that continuously selects the information to be shown to us based on our data (age, friends, location, etc.) and our previous behaviors and searches on the Internet. The consequence is that we are separated from information that disagrees with our points of view, and we all live in a different state of the Internet (= we are all in our unique universe of information). For example, on Facebook, Youtube, Twitter, Instagram, etc., even though you feel that everyone else is seeing the same thing as you, you are the only one to have this set of contents displayed to you [264].

Finally, the Internet, which gives us the impression of being with others, actually isolates us in different states of the Internet, and we end up being alone. This concept is interesting, and I encourage you to refer to the sources used [274][39][41][264][2] and this TED talk by Eli Pariser which lasts 8min48: [142].

Now that we have explained that we are enclosed in echo chambers and filter bubbles, are there any other facets of the human being to mention? Well, yes! It is worth mentioning that we tend to trust more what our community, and particularly our family and friends¹, share with us [238][164]. But also the fact that we will give more importance to articles confirming our pre-existing beliefs, attitudes, and values, and be inclined to seek information that confirms them [238][68]. So we are much more persuasive when we hear things that confirm what we think. That's why we tend to surround ourselves with people who have the same ideas as us (in Appendix B.10.1, we show you that, on Facebook, people in the U.S. have more friends on the same political side and that hatred towards the other political side is growing). We

¹As we explained in subsection 3.2.2, the experiment conducted by Facebook showed that the emotions expressed by friends, via online social networks, influence our own moods, constituting, to our knowledge, the first experimental evidence for massive-scale emotional contagion via social networks [...] [174]. Therefore, the individuals we are connected with on social networks have a real influence on our behavior.

have noticed, especially since 2016, a decrease in the confidence of mass media [238][68] (such as The New York Times, The Wall Street Journal, Le Monde, Le Figaro, ...). We have not found precise explanations, but we hypothesize that it is simply because these newspapers produce professional, objective articles that present the facts without necessarily taking a position favoring one community or another. Consequently, since people prefer articles that support their ideas, they will have less trust in neutral newspapers. Additionally, people tend to identify with personalities (such as YouTubers) and therefore accept, more easily, what they say without questioning their statements (especially among younger people).

When we combine everything that has been said, each person lives in a different world on social networks. All the information that is proposed to us has been filtered to fit us (filter bubbles) and a selective exposure to content is the primary driver of content diffusion and generates the formation of homogeneous clusters, i.e., echo chambers [127]. Therefore, we are locked in communities that share our ideas, and we are less inclined to engage in dialogue with people who hold values contrary to our own.

In summary, this entire system gives us the impression that everyone thinks like us because we are confined within these ideas. Consequently, when faced with individuals with totally opposing ideas, instead of choosing dialogue, we are more aggressive in defending our community, mainly because it may be surprising to encounter someone who does not think like everyone we see on social networks. But it is simply due to the fact that we believe that everyone thinks like that. This leads to an increase in polarization between different communities. Social networks, therefore, have a tendency to reduce tolerance for alternative views, amplify attitudinal polarization, boost the likelihood of accepting ideologically compatible news, and increase closure to new information and the dislike of the 'other side' (effective polarization) has also risen [238]. Consequently, people are more inclined to believe that fake news is true if it aligns with their existing beliefs, particularly when there is support for that information within their social circles.

5.1.1.3. Social media platforms promote sensational and viral content

Two aspects are highlighted by social networks: virality and sensationalism [186][184]. Indeed, these are two parameters that can generate a significant amount of engagement and, therefore, allow them to generate more revenue [68]. Social networks are not the only ones taking advan-

tage of this. Content creators, too, will strive to make their articles, videos, and other content as clickbait as possible. Indeed, content creators will strive to make their articles, videos, and other content as clickbait² as possible to attract the most people and generate the maximum possible revenue [186]. With the fact that anyone can share content without the required qualifications (as explained in subsection 5.1.1.1), this has increased the competition to be highlighted on social networks and attract an audience. Unfortunately, the less polarizing, sensational, and interactive the information is, the less it will be promoted because it generally generates fewer reactions. Social networks prioritize viral and sensationalist information at the expense of qualitative information [186]. This means that news about a celebrity buying a house, which generates 10,000 likes, 1 million retweets, and 100,000 comments, will be highlighted much more than news about a newly started war that only receives 10 likes... In the next section, we will see that studies have shown that fake news spreads faster and reaches more people than real news because it often generates more reactions.

People who want to spread viral information, fake news, or manipulate populations have understood this system very well. We have seen it in chapter 4, where we discussed about Cambridge Analytica and the IRA. Indeed, we observe that groups of people often create an army of bots in order to create a fake community around a subject and increase the number of reactions under their posts. They make their information more visible because it is more highlighted, they make people believe that there is a community, and finally, if the fake news is in line with a person's ideas and they see that there is support in terms of the number of reactions, they will be tempted to believe it [349][184]. Especially if the information is repeated in several posts [238].

Finally, we live in a world where the majority of people want everything instantly and for free. We want everything to go fast, to be bombarded with information, to have an explosion of emotions by reacting to everything and anything, by learning that our favorite celebrity has bought new shoes and retweeting it, to receive 1000 likes, ... that generates emotion and, above all, dopamine. It is not uncommon to see information go viral and then, after a few days or hours, disappear for another piece of information. This consumption process has become automatic: we see information and quickly retweet it to be the first and collect the maximum number of likes. This rapid flow of information prevents analysis and objective reflection on

²= Articles, photographs, etc. on the Internet that are intended to attract attention and encourage people to click on links to particular websites [1].

what we see/read. We don't take the time to check our sources and analyze the context, we want sensationalism. In addition to that, there is a very important element: the desire to have everything for free. Indeed, social networks and the Internet give us access to a wide range of quality information but also a lot of low-quality information. We conducted a small survey at UCLouvain where 181 students (mostly from the EPL and FIAL) responded (see Appendix D.1). To the question *Do you agree with the fact that you have to pay for certain articles to have access to the content?*, 67.6% (thus a majority) answered no. This means that people do not agree to pay for access to professionally written articles. They do not agree to pay journalists who have researched and verified their information, especially when there is a decrease in confidence in mass media. We have thus entered a world of hyper-consumption of information without reflection on what we consume. This clearly favors the propagation of fake news and polarization in the world. Researchers at Facebook have found that Facebook's system automatically and disproportionately spreads harmful content, and viral content favors conspiracy theories, hate speech, and hoaxes [204] (see this source [204] from The Wall Street Journal where they explain some experiments conducted by researchers at Facebook). Here is a non-exhaustive list of examples of fake news that have been amplified by social networks and have garnered a lot of support: the Pizzagate story [292][177], the false claim that Covid-19 originated from Corona beer [243][332], the false theory that Covid-19 vaccines contained microchips [194][80], the panic that took place in the U.S. with a false theory of Chinese agents saying that Trump was going to lock down the entire country [141], etc. There is also the fact that Elon Musk conducts polls on Twitter to decide whether he should unban Trump on Twitter, as if it were a game... [57].

In Appendix B.10.2, we provide you with some examples found on the Internet which, in our opinion, represent the world in which we live.

5.2. Propagation of fake news on social network during 2016 U.S. elections

We will now analyze the results of various studies that have examined the spread of fake news. While some studies agree that fake news does not seem to have had a significant impact on the 2016 U.S. election [175][176], a majority of studies agree that it did [346][186][184][127][68].

A study [175] conducted on 16,442 twitter accounts concluded that consumers of fake news were concentrated in a small portion of users and were the most active users on the network. Indeed, only 1% of panel members consumed 80.0% of the volume from fake news sources. Super-consumers of fake news tweeted an average of 71 times per day (sharing an average of 7.6 political URLs per day, 1.7 of which came from fake news sources) while the average user shared an average of 0.1 times per day. Similarly, the fake news super-consumer was exposed to nearly 4,700 political URLs daily, compared to only 49 for the median panel member. The study therefore suggests that these super consumers of fake news were bots because the difference with a normal person is very huge. They conclude that although the people sharing fake news about politics correspond to 6% of the users, the vast majority of fake news shares and exposures were attributable to tiny fractions of the population. Another study [176] conducted on Facebook (N = 3500 accounts) agrees with the previous study by explaining that sharing articles from fake news domains is a much rarer event than sharing links in general. Indeed, they found that 90% of people did not share stories from fake news domains and 8.5% of people shared at least one fake news with their friends. They also found that the people that are the most likely to share fake news are people with over 65 years and conservatives than liberals or moderates.

Now, let's look at studies claiming that fake news had an impact on the 2016 U.S. election. First, a study done by MIT [346][288] with Twitter data from 2006 to 2017 (+- 126 000 stories tweeted by +- 3 million of people more than 4,5 million times) found strong results. They sampled their data with 6 independent fact-checking organizations (snopes.com, politifact.com, factcheck.org, truthor-fiction.com, hoax-slayer.com, and urbanlegends.about.com). They found that a significantly greater fraction of false cascades³ than true cascades exceeded a depth of 10, falsehood also reached far more people than the truth. Whereas the truth rarely diffused to more than 1000 people, the top 1% of false-news cascades routinely diffused to between 1000 and 100,000 people. Falsehood reached more people at every depth of a cascade than the truth, meaning that many more people retweeted falsehood than they did the truth. The spread of falsehood was aided by its virality, meaning that falsehood did not simply spread through broadcast dynamics but rather through peer-to-peer diffusion characterized by a viral branching process. It took the truth about six times as long as falsehood to reach 1500

³A rumor cascade begins on Twitter when a user makes an assertion about a topic in a tweet, which could include written text, photos, or links to articles online. Others then propagate the rumor by retweeting it; [346]

people, etc. They also found that false news was more novel than true news, which suggests that people were more likely to share novel information. The 3 graphs that illustrate this can be found in Appendix C.1.2 (the 3 first graphs). All these conclusions can be contradicted with the argument of the presence of bot which boosts the statistics of fake news. However, they have used a sophisticated bot-detection algorithm to identify and remove all bots before running the analysis. When they added bot traffic back into the analysis, they found that none of their main conclusions changed. False news still spread farther, faster, deeper, and more broadly than the truth in all categories of information. The results remained also the same when they removed all tweet cascades started by bots, including human retweets of original bot tweets. They have also used a second independent bot detection algorithm and varied the algorithms sensitivity threshold to verify the robustness of their analysis

Another study [186] has drawn other interesting conclusions. They utilized a dataset of 22 million tweets collected between November 1 and November 11, 2016, and after sampling (retaining only tweets from individuals physically located in the U.S.), they kept 7 million tweets. Among these, 1.27 million contained URLs to websites. Firstly, they found that the number of links to Russian news stories, unverified or irrelevant links to WikiLeaks pages, or junk news was greater than the number of links to professional researched and published news. Indeed, the proportion of misinformation was twice that of the content from experts and the candidates themselves (see the figure C.4 in Appendix C.1.2). Second, a large proportion of all the successfully catalogued content provides links to polarizing content from Russian, WikiLeaks, and junk news sources. Thus, when links to Russian content and unverified WikiLeaks stories are added to the volume of junk news, 32% of all the successfully catalogued political content was polarizing, conspiracy driven, and of an untrustworthy provenance. They also did another study [184] based on the same dataset where they found that in Michigan : (1) Conversation about politics over Twitter mirrored the national trends in that Trump-related hashtags were used more than twice as often as Clinton-related hashtags. (2) They shared a lot of political content, but the amount of professionally researched political news and information was consistently smaller than the amount of extremist, sensationalist, conspiratorial, masked commentary, fake news and other forms of junk news. (3) Not only did such junk news outperform real news, but the proportion of professional news content being shared hit its lowest point the day before the election [186].

Here are some other important results found by other studies: a study [127] found that

the maximum cascade size are 952 for science news, 2422 for conspiracy news and 3945 for trolling. They also found that users mostly tend to select and share content related to a specific narrative and ignore the rest. In particular, they showed that social homogeneity is the primary driver of content diffusion, and one frequent result is the formation of homogeneous, polarized clusters. Most of the time, the information is taken by a friend who has the same profile (polarization), i.e., belonging to the same echo chamber. Another study [68] found that, according to their database, there are about three times more fake pro-Trump articles than pro-Clinton articles, and the average pro-Trump article was shared more on Facebook than the average pro-Clinton article (in their data base : 41 pro-Clinton and 115 pro-Trump articles shared, respectively 7,6 and 30,3 million times). They also found that a list of fake news websites, on which just over half of articles appear to be false, received 159 million visits during the month of the election, or 0.64 per U.S. adult.

6. Possible solutions

We have seen different issues in this thesis that are all directly or indirectly related. Few solutions exist because these are problems that are not highlighted enough and it is difficult to establish large studies due to all the restrictions (particularly at data level) we have seen.

6.1. Regulatory measures: protecting citizens' privacy and financial penalties for social media platforms

In Europe, we have the GDPR (General Data Protection Regulation) which establishes laws and rules to protect European citizens. Recently, new rules have been proposed that you can find in the following source: [213]. However, these rules are often not respected by the tech giants. For example, on May 22, 2023, Meta was fined a 'record' 1.2 billion euros [140] (= 1% of its turnover in 2022) in Europe. In the United States, in 2019, Facebook was fined 5 billion dollars by the FTC (Federal Trade Commission) [88] (= 7% of its turnover in 2019). However, even when new rules are implemented, the tech giants do not always comply with them, leading to fines. Therefore, the question is: are we being too lenient with them? 1.2 billion euros in fines represents only 1% of their turnover... So it appears that the cost/benefit ratio of engaging in fraudulent activities is in their favor. Without such fraud, their earnings may have been lower by several billion dollars. Furthermore, Europe and the USA constitute a significant portion of their user base. However, they also have a large number of users in developing countries [305] where there are no regulations in place to protect the citizens. This situation can even align with the interests of governments in those countries. Meta invests money in building infrastructure in these countries to expand its platform, but it does not invest in moderators or AI. The limited resources allocated by Meta for moderation are primarily focused on Europe and the US, where they face pressure from governments. In our

view, it is necessary to impose stricter regulations on their global operations and hold them accountable for their actions. In our view, it is imperative to adopt a more stringent approach towards these companies and establish regulations on what they do worldwide. They wanted to grow too fast because they saw that there was money to be made. However, they need to learn how to manage their platform before seeking further expansion. Additionally, in 1996, a law called *Section 230* has been adopted in the United States [342]. This law states *No provider or user of an interactive computer service shall be treated as the publisher or speaker of any information provided by another information content provider* [3]. It therefore protects these companies who do not hesitate to use it to clear themselves of any problem related to their platform [89]. It seems very complicated to destabilize them legally.

6.2. Investing in moderation tools and changing the business model

The best solution would be that they change their business model and leave this business of surveillance capitalism. But they won't do that because there is no other business model that would allow them to generate as much money and have as much power over the population with such ease. It would also require too many resources to completely change the way the platform works. Ironically, instead of investing in more moderation tools, they display prevention messages on their platform (which is not very costly in terms of money). In Appendix E.3, you will find screenshots of prevention messages that we have received during our private activities on these networks. We also have the fact that in 2018, Facebook introduced a time-management tool [168]. However, its existence seems relatively unknown as the company does not actively promote it. It would not be in their best interest if those tools were used. An article [238] suggests an intriguing solution, proposing that platforms provide users with signals of source quality that can be incorporated into the algorithmic content rankings. If done ethically this could be interesting. The question is whether these platforms would not take advantage of this ranking in exchange for money from third parties. Finally, while many Facebook employees support the implementation of moderation tools, this is not the case for the top management. In fact, a senior security official at Facebook stated that *the company would seek to disrupt on-platform movements only if there was compelling evidence that they were the product of tightly knit circles of users connected to real-world violence or other harm*

and committed to violating Facebook's rules [204]. It is quite paradoxical when considering their efforts to hinder research on Facebook and therefore prevent us from bringing evidence. Without access to all the data and algorithms, it is impossible to carry out studies that reveal the true nature of the situation [301].

6.3. Educating people on the impact of social media

Solutions will unlikely come from the tech giants themselves. Instead, they must originate from the foundation of their business model: us, the users. We, as consumers and contributors to their system, need to raise awareness about these issues that are often overlooked. It is crucial to educate [238] the younger generation, starting from secondary school, about how recommendation algorithms work and their consequences, the existence of echo chambers and filter bubbles, how to identify fake news, the effects of social media usage on our well-being (such as dopamine release and the need for social validation), the importance of source verification, the need not to react immediately to every information, the concept of information virality, and the utilization of independent fact-checking organizations (such as snopes.com, politifact.com, factcheck.org, truthorfiction.com, hoax-slayer.com, urbanlegends.about.com, etc.) [346]. Younger individuals have developing brains that are fully immersed in this system from birth, yet they have not been provided with the necessary tools to protect themselves. Therefore, it is essential to educate them on the proper utilization of these tools, as they still hold value for certain purposes. We have to reteach people to take their time, not to let themselves be manipulated by their emotions. We need to not forget that what costs us the most in this system, is the **imperceptible** consequences of the **gratuitousness**. The fact that their services are free allows them to blind us to the problem. By making us believe that everything is designed to make us happy, connect us with our loved ones, and provide us with free access to all information, they allow themselves to play with our lives behind the scenes.

6.4. Stop using social media

Finally, in my opinion, the best solution is simple and radical: do not use these social networks. That is what I have done and I feel much better. I am better informed because I have more

time to conduct my own research, I can focus better, my mental state has improved, I am more physically active, and I sleep better. I have also recently switched my smartphone screen to black and white because the colors of applications are intentionally designed to stimulate the brain. For instance, the choice of using a red notification button is not arbitrary; red creates a sense of urgency and importance. If you are interested, this video [138] talks about it and this tutorial [93] explains how to do it.

6.5. Individuals in key positions within social media companies do not use social media

We can observe that the solutions can only come from us. On one hand, we have these platforms that polarize society, and on the other hand, they addict us to the point where we no longer perceive the amount of time we dedicate to them. By getting out of this vicious circle, we will be able to gain awareness of the system in which we are immersed. The evidence lies in the fact that individuals holding important positions within these companies prohibit themselves and their children from using their own platforms [62][308][287][270]. Here is for example a sentence from Chamath Palihapitiya: *my solution is I just don't use these tools anymore. I haven't for years. [my kids] are not allowed to use this shit* [207][215].

7. Conclusion

7.1. The dark side of social networks: manipulation and control

With their business model based on surveillance capitalism, they have entered the realm of human psychology with the aim of discovering our most intimate secrets. They exploit various human vulnerabilities such as the need for belonging to a community, social validation, and more. They have meticulously designed their platforms to provide us dopamine every time. By leveraging these vulnerabilities and our addiction to dopamine, they have conditioned us to perform certain actions automatically, making us even more dependent on their platforms. They are fully aware of the harmful effects their platforms have on individuals, both intellectually and psychologically. However, driven by greed and power, they pay little attention to these concerns.

7.2. The information asymmetry and lack of transparency of social networks

They take advantage of information asymmetry and make every effort to maintain it. On one hand, this allows them to gradually and imperceptibly enter our lives. Since they know everything about us while we know nothing about them, they have the ability to manipulate us without us being aware of it. On the other hand, this enables them to prevent any studies of their platform. Scientific research is reliant on their database, which are private and contains the most valuable information. As a result, researchers must rely on public datasets that do not accurately represent reality. Furthermore, researchers are left in the dark about when and

how social media providers change the sampling/filtering of their data streams.

When we attempted to create our own database by scraping the platform and creating fake accounts, our experiences revealed that Facebook has implemented numerous tools to counter the creation of fake accounts. Even when we managed to create an account, retrieving information from the news feed became challenging. The HTML structure changes frequently, multiple XPath's are required for the same information, JSON files do not contain all the data, and various other obstacles hinder data collection. Everything seems to have been done to make data collection complex and difficult.

Additionally, when requesting access to their studies, these platforms often refuse to disclose their findings. We are only able to gain insight into their research when information is leaked or made public by other means.

7.3. Social networks facilitates propaganda, manipulation, and human trafficking

Social media platforms enable governments, companies, and third-party groups to utilize their platform and advertising targeting tools for propaganda, manipulation, incitement of hatred, human trafficking, child trafficking, and more. This has allowed Cambridge Analytica to assist several political figures, including Mr. Donald Trump, in his 2016 presidential campaign, where they manipulated the American people by obtaining the Facebook data of 80 million individuals. They also helped political parties in the UK to influence people to vote in favor of Brexit. The IRA has created numerous bots and continues to spread false information to sow chaos. They also supported Mr. Donald Trump during his 2016 presidential campaign, reaching 126 million users on Facebook, 32 million on Twitter, and receiving 185 million likes on Instagram. In Myanmar, Ethiopia, and India, there is a significant amount of hate speech inciting violence against other ethnic groups within the countries. Additionally, social media platforms are used for human trafficking purposes. Drug cartels utilize Facebook to recruit, train, and pay hitmen. Etc.

7.4. Social networks platforms prioritize fake news over authentic content

To generate ever-increasing revenue, social media platforms prioritize sensational and viral false information (which generate significant reactions) at the expense of genuine information that generates fewer reactions. In fact, the more sensational the information is, the longer users remain on the platform, enabling social media networks to generate more revenue. False information reaches 1,500 individuals six times faster than true information. Moreover, the propagation of false information is facilitated by its viral nature, indicating that it does not solely disseminate through traditional diffusion mechanisms but rather through peer-to-peer dissemination characterized by a viral branching process. Finally, a higher proportion of false cascades exceeded a depth of 10 compared to true cascades.

7.5. The inaction of social networks in addressing these issues

When it comes to the consequences of social media on individuals or societies by facilitating the spread of hatred and misinformation, Meta and other social networks platforms do not implement any measures to address these issues, despite being aware of them.

At the individual level, it suits their interests to have people become addicted to their platforms, emotionally vulnerable, and intellectually weakened. This enables them to generate increasing revenue while gathering information about users and manipulating them.

At the societal level, social media platforms generally detect content that violates their usage policies but do not take action until the media brings attention to them. This allows them to avoid wasting time and money. Such content also tends to generate reactions on their platform, thus increasing their revenue. Once the media finally addresses these issues, it is often too late, as the damage has already been done, and the platforms struggle to effectively manage the problem.

Furthermore, they heavily rely on their AI systems to detect posts that violate their usage policies, but these systems do not comprehend all languages and dialects. Additionally, two

internal teams at Facebook estimated that the AI was deleting posts that generated between 2% and 5% of the views of hate speech, as well as 0.6% of all content that violated Facebook's policies against violence and incitement. Which is not a significant amount... Moreover, there have been Facebook employees who have admitted that their information classification AI does not work effectively because it is not adapted to all cultures.

Regarding fake news, Mr. Zuckerberg has frequently emphasized the economic advantage that virality provides for his company. Moreover, when employees attempt to regulate certain features of Facebook, particularly the sharing system, top executives refuse. For instance, John Hegeman (Vice President and Head of Monetization at Meta) stated, *If we remove a small percentage of reshares from people's inventory, they decide to come back to Facebook less.*

7.6. AI guides us, the world is under its control

The recommendation algorithms of social media platforms primarily rely on artificial intelligence (AI) to suggest content to us. This AI constantly collects our data and analyzes it to determine the posts that will be displayed when we scroll through our news feed. It also contributes to the creation of filter bubbles and confines us within ideological clusters. Thus, it controls the entire content that is presented to us. It has the power to prevent us from accessing certain content or to prioritize the visibility of others. The world is therefore under its control as it decides the news topics and what should be shared. It is worth noting that at platforms like YouTube (Google) and Meta, they do not always fully understand how their AI works [264].

7.7. Social networks are at the origin of the rise of political polarization

The psychological and intellectual weakening of the population, the creation of echo chambers, filter bubbles, the faster spread of fake news compared to true information, the utilization of social media by governments for propaganda purposes, the tendency to believe what our close contacts say and articles that align with our own beliefs, and the fact that social media has

become a primary source of information for many people, all of these factors contribute to the increasing political polarization worldwide. Individuals are confined within ideological clusters where they perceive everyone to think like them. Consequently, individuals are more inclined to defend their community and ideas by asserting that they are the ones who are right. This leads to a decreased openness to debate and the emergence of animosity towards opposing ideologies. As a result, they are less open to debate and develop a hatred towards other ideologies. It has become a battle of who is right instead of taking the best aspects of each idea.

Since their inception, social media platforms have contributed to the increase in political polarization worldwide.

7.8. Future challenges: AI and Deepfakes

With the advancement of AI, particularly in the realm of deepfakes¹, the situation is bound to deteriorate. We are entering a new phase where chaos may ensue if no action is taken. It is imperative that we educate individuals and raise awareness about these issues. We must provide them with the necessary tools to comprehend the functioning of social networks and to discern between true and false information.

7.9. The potential solutions

The challenge lies in the fact that our laws evolve more slowly than technologies. This makes it difficult for our countries to effectively regulate these companies. Therefore, it is necessary to be stricter with them and not hesitate to increase fines. It should be noted that each year, Meta receive new fines, indicating that previous fines have not been sufficient to deter them from violating our rules. Additionally, in the USA, these companies are protected by the law *Section 230*.

The solution will not come from them, it must come from us. On one hand, we need to educate

¹In May 2023, there was a fabricated news report circulating about an explosion near the Pentagon. Surprisingly, this misinformation was shared by numerous verified Twitter accounts. The dissemination of this fake news had an effect on the stock market, resulting in a slight decline in certain stock prices [133].

people by explaining what is an echo chambers, a filter bubbles, what are the impact of social media on the brain, and so on. On the other hand, we can simply choose to stop using them and find more ethical alternatives for sharing information.

7.10. Final words

To provide some positive aspects of social networks: they allow content to be shared effectively (articles, discoveries, etc.) with our friends, and they have facilitated the creation of student support groups and people who share the same passion. Social networks are not entirely negative. The issue lies in their business model.

There are other topics that we did not have time to address in this thesis, such as the *Rabbit Hole* effect observed on YouTube or the models proposed by Bala and Goyal or Zollman, which can be linked to social networks [354][183][182][300][348][276]. The subject is so vast that we had to restrain ourselves to stop.

B.11.Average time spent on a smartphone in February 2021 in the U.S. for people aged 18 and older. [30]	206
B.12.Daily hours spent with digital media, United States, 2008 to 2018. [282]	207
B.13.Most visited websites in 2022 (in billion). [112]	208
B.14.Most Visited Websites by Traffic in the world for all categories, January 2023. [11]	209
B.15.Most visited website in november 2022 (in billion). [112]	210
B.16.Number of people using social media platforms, 2004 to 2018. [154]	211
B.17.Number of people using social media platforms, 2018. [282]	212
B.18.Proportion of users sharing the same items as me.	216
B.19.Attributes with how they retrieve it and an example of values used by the website https://amiunique.org/ to create a fingerprint	217
B.20.Number of days between the devices first visit and the first change of its browser fingerprint. [289]	219
B.21.Number of days between the devices first visit and the first change of its browser fingerprint depending on the type of the device. [289]	220
B.22.Structure of SCL & Cambridge Analytica Companies & Shareholders. [350]	227
B.23.Cambridge Analytica: explanation of how 50m Facebook records were hijacked. [99]	228
B.24.Cambridge Analytica: how the key players are linked. [100]	229
B.25.Worldwide web: the data companies behind Brexit. [101]	230
B.26.A campaign advert tested by Cambridge Analytica. [60]	231
B.27.A campaign advert tested by Cambridge Analytica. [60]	231
B.28.Visual messages tested by GSR’s online profiling test. Respondents were asked: How important should this message be to all Americans?. [98]	232
B.29.Visual messages tested by GSR’s online profiling test. Respondents were asked: How important should this message be to all Americans?. [98]	232
B.30.Visual messages tested by GSRs online profiling test. [98]	233
B.31.Post that spread false accusation of voter fraud in Pennsylvania. [235]	234
B.32.Post shared by the account @stay4police which sell T-shirts and hooded sweat-shirts where we can read I Support American Law Enforcement. [235]	235

B.33. Posted on March 2, 2016 by @army_of_jesus_ received nearly 90,000 likes on Instagram, making it the most-liked Instagram content uploaded by the group before the 2016 election. [284][235]	236
B.34. A post from an IRA-controlled Facebook page called Black Matters. [284] . . .	236
B.35. A post by the Kremlin’s Internet Research Agency mocking the idea of Russian interference in a post targeted to right-wing audiences. [284]	237
B.36. A post from an IRA-controlled Facebook page called Secured Borders. [302] .	237
B.37. Posted by the @blackstagram_ account in June 2017. It received more than 250,000 likes and more than 6,000 comments. [235]	238
B.38. The IRA pages offered coupons in exchange for sharing content. [132]	238
B.39. Cluster of Hillary Clinton memes, a cluster of Donald Trump memes, and other memes related to GOP politicians on the political page Being Patriotic. [132] .	239
B.40. Top 20 IRA Facebook pages sorted by number of like. [185]	240
B.41. About 6-in10 Americans get news from social media - 2016 [206]	244
B.42. Reddit, Facebook and Twitter users most likely to get news on each site - 2016 [206]	245
B.43. Social media news use: Facebook leads the pack - 2016 [206]	245
B.44. Most social media news consumers only get news on one site - 2016 [206] . . .	246
B.45. Get information by searching or by chance? - 2016 [206]	246
B.46. Growth in use of social media for news - 2016 [206]	247
B.47. Two-thirds of U.S. adults get news from social media - 2017 [205]	247
B.48. Twitter, Youtube and Snapchat grown since 2016 in portion of users who get news on each site - 2017 [205]	248
B.49. Social media sites as pathways to news - 2017 [205]	248
B.50. More Americans get news on multiple social media site - 2017 [205]	249
B.51. Americans get their information from social media, but a majority question its accuracy - 2018 [228]	250
B.52. Growth in social media news consumption slows down - 2018 [228]	250
B.53. Social media sites as pathways to news - 2018 [228]	251
B.54. Reddit, Twitter, Facebook stand out as sites with the most news-focused users - 2018 [228]	251
B.55. Most social media news consumers are concerned about inaccuracy - 2018 [228]	252
B.56. About half of Americans get news on social media at least sometimes - 2020 [70]	252

B.57. Facebook stands out as regular source of news for Americans - 2020 [70]	253
B.58. Large portion of Twitter users regularly get news on the site - 2020 [70]	254
B.59. As in past years, most social media news consumers expect news there to be inaccurate - 2020 [70]	255
B.60. About half of Americans get news on social media at least sometimes, down slightly from 2020 - 2021 [249]	255
B.61. Nearly a third of Americans regularly get news on Facebook - 2021 [249] . . .	256
B.62. Large portion of Twitter users regularly get news there - 2021 [249]	256
B.63. Homophily in self-reported ideological affiliation. [92]	258
B.64. Republicans Shift to the Right, Democrats to the Left. [255]	259
B.65. Growing Partisan Antipathy. [255]	260
B.66. U.S. hospitalisation admission for non-fatal self harm. [264]	272
B.67. U.S. suicide rates. [264]	273
B.68. Computer processing power. [264]	274
B.69. Proposition of Firefox for the sentence "the war is" (browser : Firefox, location : Belgium)	276
B.70. Proposition of Firefox for the sentence "the war is" (browser : Firefox, location : Ukraine)	277
B.71. Proposition of Edge for the sentence "the war is" (browser : Edge, location : Ukraine)	278
C.1. Complementary cumulative distribution functions (CCDFs) of false political and other types of rumor cascades. [346]	287
C.2. Complementary cumulative distribution functions (CCDFs) of true and false rumor cascades. [346]	288
C.3. Rumor cascades. [346]	288
C.4. What Political News and Information Were US Twitter Users Sharing during the Election? [186]	289
E.1. Growing minority holds consistent ideological views. [255]	305
E.2. Democrates and Republicains more ideologically dived than in the past. [255] .	306
E.3. Republicans shift to the right, Democrates to the left. [255]	307
E.4. Beyond dislike: viewing the other party as a 'Threat to the nation's well-being. [255]	308

- E.33. Title of the video: What Tinder doesn't tell you [165]. Link to the video:
https://www.youtube.com/watch?v=bxQ_RngIQBI 325
- E.34. Title of the video: 300 Days of Watching TikTok Non-Stop: WHAT DOES IT
 DO? [158]. Link to the video: https://www.youtube.com/watch?v=LdNta5n_35Q 326
- E.35. Title of the video: 50,000 Swipes on Tinder (and break his secret) [159]. Link
 to the video: <https://www.youtube.com/watch?v=9BJP2H1kZnk> 326
- E.36. Title of the video: The overproduction of YouTube [275]. Link to the video:
<https://www.youtube.com/watch?v=XgjyRVwHQQc> 327

Bibliography

- [1] clickbait, *Cambridge Dictionary*. <https://dictionary.cambridge.org/fr/dictionnaire/anglais/clickbait>. [Accessed 17.05.2022].
- [2] Filter bubble, *Wikipedia*. https://en.wikipedia.org/wiki/Filter_bubble. [Accessed 24.05.2022].
- [3] Section 230, *Electronic Frontier Foundation*. <https://www.eff.org/fr/issues/cda230>. [Accessed 23.05.2022].
- [4] Ethiopia violence: Facebook to blame, says runner Gebrselassie, *BBC*. <https://www.bbc.com/news/world-africa-50276603>, 02 November 2019. [Accessed 23.04.2023].
- [5] Dopamine : qu'est-ce que c'est ?, *Futura Science*. <https://www.futura-sciences.com/sante/definitions/medecine-dopamine-3263/>, 03 February 2019. [Accessed 04.04.2023].
- [6] Introducing Instagram Reels, *Instagram*. <https://about.instagram.com/blog/announcements/introducing-instagram-reels-announcement>, 05 August 2020. [Accessed 15.05.2022].
- [7] Qu'est-ce qu'un VPN et comment fonctionne-t-il ?, *Avast*. <https://www.avast.com/fr-fr/c-what-is-a-vpn>, 08 april 2020. [Accessed 04.10.2022].
- [8] Conditionnement opérant, *Wikipedia*. https://fr.wikipedia.org/wiki/Conditionnement_op%C3%A9rant, 10 August 2022. [Accessed 12.10.2022].
- [9] How to Avoid Bot Detection with Selenium, *ZenRows*. <https://www.zenrows.com/blog/selenium-avoid-bot-detection#how-anti-bots-work>, 11 April 2023. [Accessed 24.05.2023].
- [10] Hungary: Europe's most advertised elections, *AdLens*. <https://adlens-be>.

- medium.com/hungary-europes-most-advertised-elections-d611b4c0694c, 12 January 2022. [Accessed 18.04.2023].
- [11] Most Visited Websites by Traffic in the world for all categories, January 2023, *SemRush*. <https://fr.semrush.com/website/top/>, 13 February 2023. [Accessed 28.02.2023].
- [12] Qu'est-ce qu'un VPN, *NordVPN*. <https://nordvpn.com/fr/what-is-a-vpn/>, 13 Jun 2022. [Accessed 04.10.2022].
- [13] Dopamine : à quoi sert l'hormone du bonheur ?, *Livi*. <https://www.livi.fr/en-bonne-sante/dopamine/>, 14 November 2022. [Accessed 04.04.2023].
- [14] Sean Parker, *Forbes*. <https://www.forbes.com/profile/sean-parker/>, 15 April 2023. [Accessed 15.05.2022].
- [15] Honduras: le président sortant Hernandez réélu, *BBC*. <https://www.bbc.com/afrique/monde-42388389>, 18 December 2017. [Accessed 16.04.2023].
- [16] La Hongrie n'est plus une véritable démocratie, dénonce le Parlement européen, *Euronews*. <https://fr.euronews.com/my-europe/2022/09/15/la-hongrie-nest-plus-une-veritable-democratie-denonce-le-parlement-europeen>, 19 September 2022. [Accessed 18.04.2023].
- [17] Historical cost of computer memory and storage, *Our World in Data*. https://ourworldindata.org/grapher/historical-cost-of-computer-memory-and-storage?country=~OWID_WRL, 2022. [Accessed 24.02.2023].
- [18] NordVPN : combien d'appareils un abonnement couvre-t-il ?, *iPhonefr*. <https://www.iphon.fr/vpn/nordvpn/nombre-appareils>, 2022. [Accessed 04.10.2022].
- [19] Computational capacity of the fastest supercomputers, *Our World in Data*. <https://ourworldindata.org/grapher/supercomputer-power-flops>, 2023. [Accessed 24.02.2023].
- [20] The AdLens report on political ads on Facebook & Instagram in Belgium, 2021, *AdLens*. <https://adlens-be.medium.com/the-adlens-report-on-political-ads-on-facebook-instagram-in-belgium-2021-8b275c558>, 21 January 2022. [Accessed 18.04.2023].

- [21] éthiopie : neuf choses à savoir sur le conflit au Tigré, *Amnesty International*. <https://www.amnesty.fr/conflits-armes-et-populations/actualites/ethiopie-conflit-tigre-comprendre-interview-chercheur-amnesty-international-france> 21 Juillet 2022. [Accessed 23.04.2023].
- [22] What Is Browser Fingerprinting and How Can You Prevent It ?, *Avast*. <https://www.avast.com/c-what-is-browser-fingerprinting>, 22 October 2019. [Accessed 05.10.2022].
- [23] Qui est Evgueni Prigojine, le puissant fondateur du groupe Wagner qui ose critiquer l'armée russe, *BBC News*. <https://www.bbc.com/afrique/articles/cnkl34v18z8o>, 23 January 2023. [Accessed 09.05.2023].
- [24] Hongrie : Qui est Viktor Orban, l'ultraconservateur en route vers un quatrième mandat ?, *Euronews*. <https://fr.euronews.com/2022/03/28/hongrie-qui-est-viktor-orban-l-ultraconservateur-en-route-vers-un-quatrieme-mandat> 28 March 2022. [Accessed 18.04.2023].
- [25] Myanmar. Les systèmes de Facebook ont promu la violence contre les Rohingyas Meta doit des réparations, *Amnesty International*. <https://www.amnesty.org/fr/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations> 29 September 2022. [Accessed 21.04.2023].
- [26] En éthiopie, la guerre oubliée du Tigré, conflit le plus meurtrier du XXI^e siècle, *geo*. <https://www.geo.fr/geopolitique/en-ethiopie-la-guerre-oubliee-du-tigre-conflit-le-plus-meurtrier-du-xxie-siecle-21> 30 January 2023. [Accessed 23.04.2023].
- [27] Des utilisateurs de Facebook manipulés pour une expérience psychologique, *Le Monde*. https://www.lemonde.fr/pixels/article/2014/06/30/des-utilisateurs-de-facebook-manipules-pour-une-experience-psychologique_4447625_4408996.html, 30 Jun 2014. [Accessed 16.03.2023].
- [28] Learning in the age of the goldfish, *eHotelier*. <https://insights.ehotelier.com/announcements/2020/11/30/learning-in-the-age-of-the-goldfish/>, 30 November 2020. [Accessed 06.04.2023].

- [29] En 2022, trois quarts des humains de plus de 10 ans possèdent un téléphone portable, selon l'ONU, *LeParisien*. <https://www.leparisien.fr/high-tech/en-2022-trois-quarts-des-humains-de-plus-de-10-ans-possedent-un-telephone-portable-30-novembre-2022>. [Accessed 26.02.2023].
- [30] How much time on average do you spend on your phone on a daily basis?, *Statista*. <https://www.statista.com/statistics/1224510/time-spent-per-day-on-smartphone-us/>, February 2021. [Accessed 26.02.2023].
- [31] Market share held by leading desktop internet browsers in the United States from January 2015 to January 2023, *Statista*. <https://www.statista.com/statistics/272697/market-share-desktop-internet-browser-usa/>, February 2023. [Accessed 28.02.2023].
- [32] Browser Market Share Worldwide - January 2023, *statcounter*. <https://gs.statcounter.com/browser-market-share>, January 2023. [Accessed 28.02.2023].
- [33] Growing app complexity: Paving the way for digital lifestyles and immersive experiences, *Sandvine*. https://www.sandvine.com/hubfs/Sandvine_Redesign_2019/Downloads/2022/Phenomena%20Reports/GIPR%202022/Sandvine%20GIPR%20January%202022.pdf, January 2023. [Accessed 28.02.2023].
- [34] Part de marché du trafic Mobile vs. Desktop vs. Tablette, *SimilarWeb*. <https://www.similarweb.com/fr/platforms/>, January 2023. [Accessed 26.02.2023].
- [35] Taux d'équipement en smartphone en France en 2020, selon la tranche d'âge, *Statista*. <https://fr.statista.com/statistiques/505110/taux-de-penetration-du-smartphone-par-age-france/>, July 2021. [Accessed 26.02.2023].
- [36] Alphabet Revenue 2010-2022 — GOOG, *macrotrends*. <https://www.macrotrends.net/stocks/charts/GOOG/alphabet/revenue>, No date. [Accessed 07.04.2023].
- [37] Biography, *FoST*. <https://futureofstorytelling.org/speaker/chris-wylie>, No date. [Accessed 24.05.2022].
- [38] Bro Army, *PewDiePie Wiki*. https://pewdiepie.fandom.com/wiki/Bro_Army, No Date. [Accessed 29.05.2023].
- [39] Bulles de filtre et chambres d'écho, *Fondation Descartes*. <https://www.fondationdescartes.org/fr/actualites/bulles-de-filtre-et-chambres-d-echo>, No date. [Accessed 29.05.2023].

- fondationdescartes.org/2020/07/bulles-de-filtre-et-chambres-decho/, No date. [Accessed 24.05.2022].
- [40] Data Protection, *European Data Protection Supervisor*. https://edps.europa.eu/data-protection_en, No date. [Accessed 05.10.2022].
- [41] How Filter Bubbles Distort Reality: Everything You Need to Know, *fs*. <https://fs.blog/filter-bubbles/>, No date. [Accessed 24.05.2022].
- [42] John Hegeman, Head of Monetization, *Meta*. <https://about.meta.com/media-gallery/executives/john-hegeman/>, No date. [Accessed 21.05.2022].
- [43] Justin Rosenstein, *Forbes*. <https://www.forbes.com/profile/justin-rosenstein/>, No date. [Accessed 24.05.2022].
- [44] Langues et dialectes, *Inde en libert  *. <https://www.inde-en-liberte.com/conseils-voyage/culture/langues>, No Date. [Accessed 23.04.2023].
- [45] Moore’s law: The number of transistors on microships has doubled every two years, *Our World in Data*. <https://ourworldindata.org/uploads/2020/11/Transistor-Count-over-time.png>, No Date. [Accessed 24.02.2023].
- [46] Moore’s law: The number of transistors per microprocessor, *Our World in Data*. <https://ourworldindata.org/grapher/transistors-per-microprocessor>, No Date. [Accessed 24.02.2023].
- [47] Nucleus Accumbens & Addiction, *UK Rehab*. <https://www.uk-rehab.com/addiction/psychology/nucleus-accumbens/>, No date. [Accessed 06.04.2023].
- [48] Population & Demography Data Explorer, *Our World in Data*. <https://ourworldindata.org/explorers/population-and-demography?facet=none&Metric=Population&Sex=Both+sexes&Age+group=Total&Projection+Scenario=None&country=CHN~IND~USA~IDN~PAK>, No Date. [Accessed 24.02.2023].
- [49] Robert S. Mueller III, *WilmerHale*. <https://www.wilmerhale.com/en/people/robert-mueller>, No date. [Accessed 24.05.2022].
- [50] Roger McNamee Author of Zucked and Silicon Valley investor, *Penguin Random House speaks bureau*. <https://www.prhspeakers.com/speaker/roger-mcnamee>, No date. [Accessed 23.05.2022].

- [51] Summary of findings: An audit of Facebook’s political ad policy enforcement, *Cybersecurity For Democracy*. <https://cybersecurityfordemocracy.org/audit-facebook-political-ad-policy-enforcement>, No Date. [Accessed 26.05.2023].
- [52] Misinformation and disinformation, *American psychological association*. <https://www.apa.org/topics/journalism-facts/misinformation-disinformation>, Not date. [Accessed 18.02.2023].
- [53] presse, *LaRousse*. <https://www.larousse.fr/encyclopedie/divers/presse/82967>, Not date. [Accessed 18.02.2023].
- [54] Gunes Acar, Marc Juarez, Nick Nikiforakis, Claudia Diaz, Seda Gürses, Frank Piessens, and Bart Preneel. Fpdetector: dusting the web for fingerprinters. In *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security*, pages 1129–1140, 2013.
- [55] Ponticelli Adèle and Grynszpan Emmanuel. Groupe Wagner : Evgueni Prigojine, une menace pour Vladimir Poutine ?, *Le Monde*. https://www.lemonde.fr/podcasts/article/2023/04/03/groupe-wagner-evgueni-prigojine-une-menace-pour-vladimir-poutine_6168017_5463015.html, 03 April 2023. [Accessed 09.05.2023].
- [56] Chen Adrian. The Agency, *The New York Times*. <https://www.nytimes.com/2015/06/07/magazine/the-agency.html>, 02 June 2015. [Accessed 24.04.2023].
- [57] AFP. Réintégrer Trump sur Twitter: le oui est majoritaire dans le sondage d’Elon Musk, *Le Soir*. <https://www.lesoir.be/478037/article/2022-11-19/reintegrer-trump-sur-twitter-le-oui-est-majoritaire-dans-le-sondage-delon-musk>, 19 September 2022. [Accessed 17.05.2022].
- [58] Le Figaro & AFP. Washington qualifie de génocide les violences de l’armée birmane contre les Rohingyas, *Le Figaro*. <https://www.lefigaro.fr/flash-actu/washington-qualifie-de-genocide-les-violences-de-l-armee-birmane-contre-les-rohing>, 21 March 2022. [Accessed 21.04.2023].
- [59] Guichard Albane. Pourquoi Facebook a-t-il racheté Whatsapp ? Instagram : de sa création à son rachat par Facebook, 8 anecdotes que vous ig-

- norez sûrement, *Capital*. <https://www.capital.fr/entreprises-marches/instagram-de-sa-creation-a-son-rachat-par-facebook-8-anecdotes-que-vous-ignorez-su> 15 February 2022. [Accessed 04.03.2023].
- [60] Hern Alex. Cambridge Analytica: how did it turn clicks into votes?, *The Guardian*. <https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie>, 06 May 2018. [Accessed 09.05.2023].
- [61] Hern Alex. Facebook owner Meta to sack 11,000 workers after revenue collapse, *The Guardian*. <https://www.theguardian.com/technology/2022/nov/09/mark-zuckerberg-meta-to-sack-11000-workers-after-revenue-collapse-facebook-instagram> 09 November 2022. [Accessed 17.05.2022].
- [62] Hern Alex. 'Never get high on your own supply' why social media bosses don't use social media, *The Guardian*. <https://www.theguardian.com/media/2018/jan/23/never-get-high-on-your-own-supply-why-social-media-bosses-dont-use-social-media>, 23 January 2018. [Accessed 09.10.2022].
- [63] Hern Alex and Cadwalladr Carole. Revealed: Aleksandr Kogan collected Facebook users' direct messages, *The Guardian*. <https://www.theguardian.com/uk-news/2018/apr/13/revealed-aleksandr-kogan-collected-facebook-users-direct-messages>, 13 April 2018. [Accessed 07.04.2023].
- [64] York Alex. Portée vs. impressions : quelle est la différence entre ces termes ?, *Sprout-Social*. https://sproutsocial.com/fr/insights/reach-vs-impressions-fr_fr/, 7 August 2020. [Accessed 26.05.2023].
- [65] Nix Alexander. How big data got the better of Donald Trump, *campaign*. <https://www.campaignlive.co.uk/article/big-data-better-donald-trump/1383025>, 10 February 2016. [Accessed 14.04.2023].
- [66] Samur Alexandra. The History of Social Media: 29+ Key Moments, *Hootsuite*. <https://blog.hootsuite.com/history-social-media/>, 22 November 2018. [Accessed 04.03.2023].
- [67] Orsini Alexis. Facebook : la désinformation russe a provoqué des rassemblements anti-immigration aux USA, *Numerama*. <https://www.numerama.com/politique/>

- 288375-facebook-la-desinformation-russe-a-provoque-des-rassemblements-anti-immigra.html, 12 September 2017. [Accessed 28.04.2023].
- [68] Hunt Allcott and Matthew Gentzkow. Social media and fake news in the 2016 election. *Journal of economic perspectives*, 31(2):211–236, 2017.
- [69] Taub Amanda and Fisher Max. How YouTube Misinformation Resolved a WhatsApp Mystery in Brazil, *The New York Times*. <https://www.nytimes.com/2019/08/15/the-weekly/how-youtube-misinformation-resolved-a-whatsapp-mystery-in-brazil.html>, 15 August 2019. [Accessed 23.04.2023].
- [70] Mitchell Amy and Shearer Elisa. News Use Across Social Media Platforms in 2020, *Pew Research Center*. <https://www.pewresearch.org/journalism/2021/01/12/news-use-across-social-media-platforms-in-2020/>, 12 January 2021. [Accessed 18.05.2022].
- [71] Benabs Ana. Intermittent Reinforcement, *AlleyDog*. <https://www.alleydog.com/glossary/definition.php?term=Intermittent+Reinforcement>. [Accessed 15.05.2022].
- [72] Benabs Ana. Sean Parker, l’ancien président de Facebook, balance sur les dangers du réseau social, *France24*. <https://www.france24.com/fr/20171110-sean-parker-lancien-president-facebook-balance-dangers-reseau-social>, 10 November 2017. [Accessed 15.05.2022].
- [73] Lucina Anaëlle. Un humain passe en moyenne 4,8 heures par jour sur son smartphone, *Le Soir*. <https://geeko.lesoir.be/2022/01/17/un-humain-passe-en-moyenne-48-heures-par-jour-sur-son-smartphone/>, 17 January 2022. [Accessed 26.02.2023].
- [74] Downie Andrew. YouTube and Facebook letting Brazil election disinformation spread, NGO says, *The Guardian*. <https://www.theguardian.com/world/2022/oct/20/brazil-election-disinformation-youtube-facebook>, 20 October 2022. [Accessed 23.04.2023].
- [75] Sinha Roy Ankush. How does facebook handle the 4+ petabyte of data generated per day? Cambridge Analytica - face-

- book data scandal., *Medium*. <https://medium.com/@srank2000/how-facebook-handles-the-4-petabyte-of-data-generated-per-day-ab86877956f4>, 16 September 2020. [Accessed 24.02.2023].
- [76] Lembke Anna. About, *Anna Lembked, MD*. <https://www.annalembke.com/about>, 2021. [Accessed 26.05.2023].
- [77] Della Valle Anne-Christine. Dopamine : à quoi sert cette 'hormone du bonheur' ?, *Le journal des femmes*. <https://sante.journaldesfemmes.fr/fiches-anatomie-et-examens/2519980-dopamine-hormone-definition-fonction-sport-drogues/>, 17 october 2022. [Accessed 04.04.2023].
- [78] Boadle Anthony. Facebook's WhatsApp flooded with fake news in Brazil election, *Reuters*. <https://www.reuters.com/article/us-brazil-election-whatsapp-explainer-idUSKCN1MU0UP>, 20 October 2018. [Accessed 23.04.2023].
- [79] Troianovski Anton. A former Russian troll speaks: 'It was like being in Orwell's world', *The Washington Post*. <https://www.washingtonpost.com/news/worldviews/wp/2018/02/17/a-former-russian-troll-speaks-it-was-like-being-in-orwells-world/>, 17 February 2018. [Accessed 24.04.2023].
- [80] Maad Assma and Audureau William. Il aurait financé la pandémie, planifié la vaccination... Ce que vous avez lu sur Bill Gates est-il vrai ?, *Le Monde*. https://www.lemonde.fr/les-decodeurs/article/2020/06/23/oms-vaccins-puce-sous-cutanee-ce-que-vous-avez-lu-sur-bill-gates-est-il-vrai_6043910_4355770.html, 24 Jun 2020. [Accessed 28.10.2022].
- [81] Touriel Aubry. En Belgique, quels partis politiques ont dépensé le plus en publicités Facebook en 2020 ?, *RTBF*. <https://www.rtbf.be/article/en-belgique-quels-partis-politiques-ont-depense-le-plus-en-publicites-facebook-en-> 09 December 2021. [Accessed 18.04.2023].
- [82] Touriel Aubry. La N-VA, le Vlaams Belang et le PTB sont les partis qui dépensent le plus en pub sur Facebook en 2022, *RTBF*. <https://www.rtbf.be/article/la-n-va-le-vlaams-belang-et-le-ptb-sont-les-partis-qui-depensent-le-plus-en-pub-s>

- 11 January 2023. [Accessed 18.04.2023].
- [83] Touriel Aubry. Le Vlaams Belang lance son propre média, les autres partis vont-ils suivre ?, *RTBF*. <https://www.rtbef.be/article/le-vlaams-belang-lance-sa-propre-chaine-les-autres-partis-vont-ils-suivre-10635543>, 19 November 2021. [Accessed 18.04.2023].
- [84] Tips Audrey. Audience similaire, *Audreytips*. <https://audreytips.com/glossaire-web/audience-similaire-facebook/>, 28 Novembre 2020. [Accessed 11.04.2023].
- [85] Aurore. Quelles sont les religions majoritaires au Myanmar, *Birmanie*. <https://www.birmanie.net/religions-birmanie/>, 05 April 2019. [Accessed 21.04.2023].
- [86] Le Monde avec AFP. Après deux ans d’une guerre sanglante en Ethiopie, gouvernement et rebelles du Tigré signent un accord pour mettre fin aux hostilités, *Le Monde*. https://www.lemonde.fr/afrique/article/2022/11/02/guerre-en-ethiopie-apres-deux-ans-de-conflit-au-tigre-rebelles-et-forces-gouvernement-6148258_3212.html, 02 November 2022. [Accessed 23.04.2023].
- [87] Le Monde avec AFP. Hongrie : victorieux des législatives, Viktor Orban se dirige vers un quatrième mandat successif, *Le Monde*. https://www.lemonde.fr/international/article/2022/04/03/hongrie-viktor-orban-revendique-la-victoire-aux-legislatives-et-se-dirige-vers-un-6120413_3210.html, 03 April 2022. [Accessed 18.04.2023].
- [88] Le Monde avec AFP. Facebook : l’amende record de 5 milliards de dollars validée par un juge, *Le Monde*. https://www.lemonde.fr/pixels/article/2020/04/25/facebook-l-amende-record-de-cinq-milliards-de-dollars-validee-par-un-juge_6037778_4408996.html, 25 April 2020. [Accessed 23.05.2022].
- [89] L’EXPRESS.fr avec AFP. Les géants du Net devant la Cour suprême : cette ‘section 230’ qui les protège, *L’express*. <https://www.lexpress.fr/economie/high-tech/les-geants-du-net-devant-la-cour-supreme-tout-savoir-sur-la-section-230-qui-les-protège>, 22 February 2023. [Accessed 23.05.2022].
- [90] Perera Ayesh. Fluid and Crystallized Intelligence (Cattell), *Simply Psychology*. <https://www.simplypsychology.org/fluid-crystallized-intelligence.html>, 08 February 2023.

- ary 2023. [Accessed 20.05.2022].
- [91] Babbar. Milliardaire du web : Sean Parker, *Babbar*. <https://babbar.academy/milliardaire-du-web-sean-parker/>, 08 September 2022. [Accessed 15.05.2022].
- [92] Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015.
- [93] Krasnoff Barbara. How to turn your phone screen to grayscale, *The Verge*. <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/>, 14 March 2023. [Accessed 21.05.2022].
- [94] L. Bastien. Pourquoi la Data Visualization aide le cerveau comprendre les informations, *LeBigData.fr*. <https://www.lebigdata.fr/pourquoi-datavis-cerveau>, 02 April 2021. [Accessed 31.05.2023].
- [95] Benot Baudry and Pierre Laperdrix. Le fingerprinting : une nouvelle technique de tra  age, *Ed-Diamond*. <https://connect.ed-diamond.com/MISC/misc-081/le-fingerprinting-une-nouvelle-technique-de-tracage>, September 2015. [Accessed 03.10.2022].
- [96] Robert M Bond, Christopher J Fariss, Jason J Jones, Adam DI Kramer, Cameron Marlow, Jaime E Settle, and James H Fowler. A 61-million-person experiment in social influence and political mobilization. *Nature*, 489(7415):295–298, 2012.
- [97] Kutuchief Britny. Reels Instagram : guide simplifi      destination des entreprises, *Hootsuite*. <https://blog.hootsuite.com/fr/reels-instagram-guide/>, 26 October 2022. [Accessed 15.05.2022].
- [98] Cadwalladr Carole. 'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower, *The Guardian*. <https://www.theguardian.com/news/2018/mar/17/data-war-whistleblower-christopher-wylie-faceook-nix-bannon-trump>, 18 March 2018. [Accessed 14.04.2023].
- [99] Cadwalladr Carole and Graham-Harrison Emma. How Cambridge Analytica turned Facebook 'likes' into a lucrative political tool, *The Guardian*. <https://www.theguardian.com/technology/2018/mar/17/facebook-cambridge-analytica-kogan-data-algorithm>, 17 March 2018. [Ac-

- cessed 07.04.2023].
- [100] Cadwalladr Carole and Graham-Harrison Emma. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach, *The Guardian*. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>, 17 March 2018. [Accessed 07.04.2023].
- [101] Cadwalladr Carole and Townsend Mark. Revealed: the ties that bound Vote Leave’s data firm to controversial Cambridge Analytica, *The Guardian*. <https://www.theguardian.com/uk-news/2018/mar/24/aggregateiq-data-firm-link-raises-leave-group-questions>, 24 March 2018. [Accessed 14.04.2023].
- [102] Allen Caroline. Facebook’s Content Moderation Failures in Ethiopia, *Council on Foreign Relations*. <https://www.cfr.org/blog/facebooks-content-moderation-failures-ethiopia>, 19 April 2022. [Accessed 22.04.2023].
- [103] Newton Casey. The tier list: how Facebook decides which countries need protection, *The Verge*. <https://www.theverge.com/22743753/facebook-tier-list-countries-leaked-documents-content-moderation>, 25 October 2021. [Accessed 22.04.2023].
- [104] Ketfi Cassim. En France, plus de 50 % du trafic internet passe par les géants américains (Netflix, Google, Amazon, Facebook), *Frandroid*. https://www.frandroid.com/culture-tech/web/1386317_en-france-plus-de-50-du-traffic-internet-passe-par-les-geants-americains-netflix-g 01 July 2022. [Accessed 28.02.2023].
- [105] Zakrzewski Cat, De Vynck Gerrit, Masih Niha, and Shibani Mahtani. How Facebook neglected the rest of the world, fueling hate speech and violence in India, *The Washington Post*. <https://www.washingtonpost.com/technology/2021/10/24/india-facebook-misinformation-hate-speech/>, 24 October 2021. [Accessed 23.04.2023].
- [106] Arthur Charles. Facebook’s mobile journey has only just begun, but already makes money, *TheGuardian*. <https://www.theguardian.com/technology/2014/feb/03/>

- facebook-mobile-desktop-pc-platforms, 03 February 2014. [Accessed 04.03.2023].
- [107] Duhigg Charles. How Companies Learn Your Secrets, *The New York Times Magazine*. <https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html>, 16 February 2012. [Accessed 18.03.2023].
- [108] Woitier Chloé. Facebook lorgne l'un des principaux opérateurs télécom d'Inde, Reliance Jio, *Le Figaro*. <https://www.lefigaro.fr/secteur/high-tech/facebook-lorgne-l-un-des-principaux-operateurs-telecom-d-inde-reliance-jio-20200324> 24 March 2020. [Accessed 23.04.2023].
- [109] Vallance Chris. Son sues Meta over father's killing in Ethiopia, *BBC*. <https://www.bbc.com/news/technology-63938628>, 14 December 2022. [Accessed 23.04.2023].
- [110] Bonnington Christina. Here's how Facebook tracks your mouse movements onscreen, *daily dot*. <https://www.dailydot.com/debug/facebook-mouse-movements/>, 14 Jun 2018. [Accessed 23.05.2022].
- [111] Asselin Christophe. Facebook les chiffres essentiels en 2023 en France et dans le Monde, *DigitMind*. <https://blog.digimind.com/fr/agences/facebook-chiffres-essentiels>, January 2023. [Accessed 04.03.2023].
- [112] Villiers Claire. Les sites web les plus visités, *Statista*. <https://fr.statista.com/infographie/28488/sites-internet-les-plus-consultes-dans-le-monde/>, 22 November 2022. [Accessed 28.02.2023].
- [113] Duffy Clare. Facebook has known it has a human trafficking problem for years. It still hasn't fully fixed it, *CNN Business*. <https://edition.cnn.com/2021/10/25/tech/facebook-instagram-app-store-ban-human-trafficking/index.html>, 25 October 2021. [Accessed 17.05.2022].
- [114] Fraser Colin. Target didn't figure out a teenager was pregnant before her father did, and that one article that said they did was silly and bad., *Medium*. <https://medium.com/@colin.fraser/target-didnt-figure-out-a-teen-girl-was-pregnant-before-her-father-did-a6be13b973a> 04 January 2020. [Accessed 18.03.2023].
- [115] Zoe Corbyn. Facebook experiment boosts us voter turnout. *Nature*, 12, 2012.
- [116] Timberg Craig. Russian propaganda effort helped spread

- 'fake news' during election, experts say, *The Washington Post*. https://www.washingtonpost.com/business/economy/russian-propaganda-effort-helped-spread-fake-news-during-election-experts-say/2016/11/24/793903b6-8a40-4ca9-b712-716af66098fe_story.html, 24 November 2016. [Accessed 23.05.2022].
- [117] Tardguila Cristina, Benevenuto Fabrcio, and Pablo Ortellado. Fake News Is Poisoning Brazilian Politics. WhatsApp Can Stop It., *The New York Times*. <https://www.nytimes.com/2018/10/17/opinion/brazil-election-fake-news-whatsapp.html>, 17 October 2018. [Accessed 23.04.2023].
- [118] Louis Cyrille. Italie : Fratelli d'Italia, de Giorgia Meloni, remporte les   lections g  n  rales, *Le Figaro*. <https://www.lefigaro.fr/international/italie-fratelli-d-italia-de-giorgia-meloni-remporte-les-elections-generales-20220906>, 06 September 2022. [Accessed 18.04.2023].
- [119] Milmo Dan. Elon Musk fires additional 200 people at Twitter, report says, *The Guardian*. <https://www.theguardian.com/technology/2023/feb/27/elon-musk-fires-additional-200-people-at-twitter-report-says>, 27 February 2023. [Accessed 17.05.2022].
- [120] Avelar Daniel. WhatsApp fake news during Brazil election 'favoured Bolsonaro', *The Guardian*. <https://www.theguardian.com/world/2019/oct/30/whatsapp-fake-news-brazil-election-favoured-jair-bolsonaro-analysis-suggests>, 30 October 2019. [Accessed 23.04.2023].
- [121] Dib Daniela and Froio Nicole. Brazilian authorities step in to catch fake news faster than Facebook, *Rest oO World*. <https://restofworld.org/2022/brazil-fake-news/>, 28 October 2022. [Accessed 23.04.2023].
- [122] Gregorian Dareh. Who is Robert Mueller, the man behind the report on Trump?, *NBC News*. <https://www.nbcnews.com/politics/justice-department/who-robert-mueller-man-behind-report-trump-n974296>, 23 March 2019. [Accessed 24.05.2022].
- [123] Alba Davey. How Russia's Troll Farm Is Changing Tactics Before the Fall Election, *The New York Times*. <https://www.nytimes.com/2020/03/29/technology/russia-troll-farm-election.html>, 29 March 2020. [Accessed 16.05.2022].

- [124] Ingram David. Factbox: Who is Cambridge Analytica and what did it do?, *Reuters*. <https://www.reuters.com/article/us-facebook-cambridge-analytica-factbox-idUSKBN1GW07F>, 20 March 2018. [Accessed 20.05.2022].
- [125] Smith David. Mueller’s indictment of 13 Russians strikes at the heart of the meddling matter, *The Guardian*. <https://www.theguardian.com/us-news/2018/feb/16/trump-mueller-russia-indictment-what-does-it-mean>, 16 February 2018. [Accessed 28.04.2023].
- [126] Seetharaman Deepa, Horwitz Jeff, and Scheck Justin. Facebook Says AI Will Clean Up the Platform. Its Own Engineers Have Doubts., *The Wall Street Journal*. https://www.wsj.com/articles/facebook-ai-enforce-rules-engineers-doubtful-artificial-intelligence-11634338184?mod=article_inline, 17 October 2021. [Accessed 21.05.2022].
- [127] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the national academy of Sciences*, 113(3):554–559, 2016.
- [128] Le Feuvre Delphine. D  couvrez le classement des 8 langues les plus parl  es au monde en 2023, *Geo*. <https://www.geo.fr/voyage/quelles-sont-les-langues-les-plus-parlees-au-monde-203079>, 20 April 2023. [Accessed 23.04.2023].
- [129] Georgiev Deyan. How Much Time Does the Average American Spend on Their Phone in 2023?, *techjury*. <https://techjury.net/blog/how-much-time-does-the-average-american-spend-on-their-phone/>, 12 January 2023. [Accessed 26.02.2023].
- [130] Georgiev Deyan. How Much Time Do People Spend on Social Media in 2023?, *techjury*. <https://techjury.net/blog/time-spent-on-social-media/>, 28 February 2023. [Accessed 28.02.2023].
- [131] Pilipenko Diana and Dessel Talia. Following the Money: Trump and Russia-Linked Transactions From the Campaign to the Presidential Inauguration, *CAP*. <https://www.americanprogress.org/article/following-the-money/>, 17 December 2018.

- [Accessed 16.05.2022].
- [132] Renee DiResta, Kris Shaffer, Becky Ruppel, David Sullivan, Robert Matney, Ryan Fox, Jonathan Albright, and Ben Johnson. The tactics & tropes of the internet research agency. 2019.
- [133] O’Sullivan Donie. Verified Twitter accounts share fake image of ‘explosion’ near Pentagon, causing confusion, *CNN Business*. <https://edition.cnn.com/2023/05/22/tech/twitter-fake-image-pentagon-explosion/index.html>, 23 May 2023. [Accessed 26.05.2023].
- [134] O’Sullivan Donie, Subramaniam Tara, and Duffy Clare. Not stopping ‘Stop the Steal:’ Facebook Papers paint damning picture of company’s role in insurrection, *CNN Business*. <https://edition.cnn.com/2021/10/22/business/january-6-insurrection-facebook-papers/index.html>, 24 October 2021. [Accessed 17.05.2022].
- [135] Kristianto Donny. Winning the Attention War: Consumers in Nine Major Markets Now Spend More than Four Hours a Day in Apps, *Data.ai*. <https://www.data.ai/en/insights/market-data/q1-2021-market-index>, 8 April 2021. [Accessed 28.02.2023].
- [136] Léo Duff. Comment TikTok a ba*sé le cerveau d’une génération., *YouTube*. <https://www.youtube.com/watch?v=FALDuZ5pPxc>, 05 February 2023. [Accessed 22.05.2023].
- [137] Léo Duff. Le système pour réinitialiser votre cerveau., *YouTube*. <https://www.youtube.com/watch?v=eUB6zkW-e5Y>, 15 February 2023. [Accessed 22.05.2023].
- [138] Léo Duff. Ce réglage peut briser votre addiction au smartphone., *YouTube*. <https://www.youtube.com/watch?v=U0JUgXzC2CI&t=1s>, 17 May 2023. [Accessed 22.05.2023].
- [139] Byers Dylan. Facebook estimates 126 million people were served content from Russia-linked pages, *CNN Business*. <https://money.cnn.com/2017/10/30/media/russia-facebook-126-million-users/index.html>, 31 October 2017. [Accessed 28.04.2023].
- [140] EDPB. 1.2 billion euro fine for Facebook as a result of EDPB binding decision, *EDPB*. https://edpb.europa.eu/news/news/2023/12-billion-euro-fine-facebook-result-edpb-binding-decision_en, 22 May 2023. [Accessed 21.05.2022].

- [141] Wong Edward, Rosenberg Matthew, and Barnes Julia E. Chinese Agents Helped Spread Messages That Sowed Virus Panic in U.S., Officials Say, *The New York Times*. <https://www.nytimes.com/2020/04/22/us/politics/coronavirus-china-disinformation.html>, 22 April 2020. [Accessed 16.10.2022].
- [142] Pariser Eli. Beware online "filter bubbles", *TED*. https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles, 2011. [Accessed 24.05.2022].
- [143] Braun Elisa. Google, Facebook et Twitter dans la tourmente de l'enqu  te russe, *Le Figaro*. <https://www.lefigaro.fr/secteur/high-tech/2017/10/11/32001-20171011ARTFIG00271-google-facebook-et-twitter-dans-la-tourmente-de-l-enquet.php>, 11 October 2017. [Accessed 28.04.2023].
- [144] Mackintosh Eliza. Facebook knew it was being used to incite violence in Ethiopia. It did little to stop the spread, documents show, *BBC Business*. <https://edition.cnn.com/2021/10/25/business/ethiopia-violence-facebook-papers-cmd-intl/index.html>, 25 October 2021. [Accessed 23.04.2023].
- [145] Dwoskin Elizabeth, Newmyer Tory, and Mahtani Shibani. The case against Mark Zuckerberg: Insiders say Facebook's CEO chose, *The Washington Post*. <https://www.washingtonpost.com/technology/2021/10/25/mark-zuckerberg-facebook-whistleblower/>, 25 October 2021. [Accessed 23.04.2023].
- [146] Barry Ellen. Long Before Cambridge Analytica, a Belief in the 'Power of the Subliminal', *The New York Times*. <https://www.nytimes.com/2018/04/20/world/europe/oakes-scl-cambridge-analytica-trump.html>, 20 April 2018. [Accessed 14.04.2023].
- [147] Schrage Elliot. Hard Questions: Russian Ads Delivered to Congress, *Meta*. <https://about.fb.com/news/2017/10/hard-questions-russian-ads-delivered-to-congress/>, 02 October 2017. [Accessed 28.04.2023].
- [148] Gilmer Emilie. Pourquoi finit-on par voter comme nos parents ?, *Femme Actuelle*. <https://www.femmeactuelle.fr/sante/psycho/elections-vote-familial-education-politique-2039150>, 28 February 2017. [Accessed 19.05.2022].

- [149] Turrettini Emily. La technologie persuasive nous rend accros aux écrans, *Le Temps*. <https://www.letemps.ch/opinions/technologie-persuasive-rend-accros-aux-ecrans>, 20 november 2016. [Accessed 12.10.2022].
- [150] Woollacott Emma. Russian Trolls Outsource Disinformation Campaigns To Africa, *Forbes*. <https://www.forbes.com/sites/emmawoollacott/2020/03/13/russian-trolls-outsource-disinformation-campaigns-to-to-africa/>, 13 March 2020. [Accessed 24.04.2023].
- [151] Akinwotu Emmanuel. Facebook’s role in Myanmar and Ethiopia under new scrutiny, *The Guardian*. <https://www.theguardian.com/technology/2021/oct/07/facebooks-role-in-myanmar-and-ethiopia-under-new-scrutiny>, 07 October 2021. [Accessed 21.04.2023].
- [152] le Boulout Eric and Zanchi Jean-Sébastien. Les 15 dates qui ont fait le Web, *01Net*. <https://www.01net.com/actualites/les-15-dates-qui-ont-fait-le-web-615826.html>, 12 March 2019. [Accessed 24.02.2023].
- [153] Siegel Eric. When Algorithms Infer Pregnancy or Other Sensitive Information About People, *MAIEI*. <https://montrealethics.ai/when-algorithms-infer-pregnancy-or-other-sensitive-information-about-people/>, 02 November 2020. [Accessed 18.03.2023].
- [154] Ortiz-Ospina Esteban. The rise of social media, *Our World in Data*. <https://ourworldindata.org/rise-of-social-media>, 18 September 2019. [Accessed 04.03.2023].
- [155] Nguyen Estelle. Propagande russe pro-Trump : 126 millions d’Américains touchés sur Facebook, *La Tribune*. <https://www.latribune.fr/technos-medias/internet/propagande-russe-pro-trump-126-millions-d-americaains-touchees-sur-facebook-801440.html>, 18 October 2018. [Accessed 28.04.2023].
- [156] Le Figaro et Afp et reuters Agences. Facebook a conduit une expérience secrète sur 700.000 utilisateurs, *Le Figaro*. <https://www.lefigaro.fr/secteur/high-tech/2014/06/30/32001-20140630ARTFIG00096-facebook-a-conduit-une-experience-secrete-sur-700000-uti>

- php, 30 June 2014. [Accessed 15.05.2022].
- [157] Randanne Fabien. France Télévisions: Fini les pubs dans les programmes pour enfants!, *20 Minutes*. <https://www.20minutes.fr/television/1976839-20161208-france-televisions-fini-pubs-programmes-enfants>, 08 December 2016. [Accessed 21.05.2022].
- [158] Poisson Fécond. 300 jours à Regarder TikTok Non-Stop : ÇA FAIT QUOI ?, *YouTube*. https://www.youtube.com/watch?v=LdNta5n_35Q, 13 May 2023. [Accessed 22.05.2023].
- [159] Poisson Fécond. 50 000 Swipes sur Tinder (et percer son secret), *YouTube*. <https://www.youtube.com/watch?v=9BJP2H1kZnk>, 29 April 2023. [Accessed 22.05.2023].
- [160] Brian J Fogg. Persuasive computers: perspectives and research directions. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 225–232, 1998.
- [161] Brian J Fogg. Persuasive technology: using computers to change what we think and do. *Ubiquity*, 2002(December):2, 2002.
- [162] Agence France-Presse. Le président sortant se déclare vainqueur au Honduras, *LeDevoir*. <https://www.ledevoir.com/monde/ameriques/514054/honduras-le-president-sortant-se-declare-vainqueur-avant-les-resultats-officiels>, 27 November 2017. [Accessed 16.04.2023].
- [163] Dibouës Franck. Sérotonine versus dopamine, *Franck Dibouës*. <https://www.psychologuesingapour.com/serotonine-versus-dopamine/>, 27 February 2020. [Accessed 04.04.2023].
- [164] Adrien Friggeri, Lada Adamic, Dean Eckles, and Justin Cheng. Rumor cascades. In *proceedings of the international AAAI conference on web and social media*, volume 8, pages 101–110, 2014.
- [165] Gaspard G. Ce que tinder vous cache, *YouTube*. https://www.youtube.com/watch?v=bxQ_RngIQBI, 13 May 2023. [Accessed 22.05.2023].
- [166] Cuofano Gennaro. How much money does Google make from advertising?, *FourWeekMBA*. <https://fourweekmba.com/how-much-money-does-google-make-from-advertising/>. [Accessed 07.04.2023].

- [167] Cuofano Gennaro. Tout savoir sur les Reels Instagram, *FourWeekMBA*. <https://fourweekmba.com/fr/ce-qui-est-arriv%C3%A9-%C3%A0-sean-parker/>, 03 April 2023. [Accessed 15.05.2022].
- [168] Wells Georgia, Seetharaman Deepa, and Horwitz Jeff. How TikTok Serves Up Sex and Drug Videos to MinorsIs Facebook Bad for You? It Is for About 360 Million Users, Company Surveys Suggest, *The Wall Street Journal*. https://www.wsj.com/articles/facebook-bad-for-you-360-million-users-say-yes-company-documents-facebook-files-11632849667?mod=article_inline, 05 November 2021. [Accessed 21.05.2022].
- [169] Wells Georgia and Horwitz Jeff. Facebook’s Effort to Attract Preteens Goes Beyond Instagram Kids, Documents Show, *The Wall Street Journal*. https://www.wsj.com/articles/facebook-instagram-kids-tweens-attract-11632849667?mod=article_inline, 28 September 2021. [Accessed 21.05.2022].
- [170] Wells Georgia, Horwitz Jeff, and Seetharaman Deepa. Facebook Knows Instagram Is Toxic for Teen Girls, Company Documents Show, *The Wall Street Journal*. https://www.wsj.com/articles/facebook-knows-instagram-is-toxic-for-teen-girls-company-documents-show-1163162073?mod=hp_lead_pos7&mod=article_inline, 14 September 2021. [Accessed 21.05.2022].
- [171] Wright Georgina and Vidotto Labastie Cécilia. L’Institut Montaigne éclaire : les élections italiennes, *Institut Montaigne*. <https://www.institutmontaigne.org/expressions/linstitut-montaigne-eclaire-les-elections-italiennes>, 20 October 2022. [Accessed 18.04.2023].
- [172] Delacroix Géraldine. Roger McNamee: Personne ne devrait être autorisé à vous suivre sur Internet, *Mediapart*. <https://www.mediapart.fr/journal/economie/290919/roger-mcnamee-personne-ne-devrait-etre-autorise-vous-suivre-sur-internet>, 29 September 2019. [Accessed 23.05.2022].
- [173] Polito Gianni. La petite histoire du web mobile, *Medium*. <https://giannipolito.medium.com/la-petite-histoire-du-web-mobile-d070f7273b98>, 06 September 2017. [Accessed 26.02.2023].
- [174] McNeal Gregory S. Facebook Manipulated User News Feeds To Create Emotional Responses, *Forbes*. <https://www.forbes.com/sites/gregorymcneal/2014/06/28/facebook-manipulated-user-news-feeds-to-create-emotional-contagion/>, 28

- June 2014. [Accessed 15.05.2022].
- [175] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. Fake news on twitter during the 2016 us presidential election. *Science*, 363(6425):374–378, 2019.
- [176] Andrew Guess, Jonathan Nagler, and Joshua Tucker. Less than you think: Prevalence and predictors of fake news dissemination on facebook. *Science advances*, 5(1):eaau4586, 2019.
- [177] Andrew Guess, Brendan Nyhan, and Jason Reifler. Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 us presidential campaign. 2018.
- [178] Champeau Guillaume. Facebook a testé sa capacité de manipulation mentale des foules, *Numerama*. <https://www.numerama.com/sciences/29852-facebook-a-teste-sa-capacite-de-manipulation-mentale-des-foules.html>, 30 Jun 2014. [Accessed 16.03.2023].
- [179] Ente Gwendoline. 1997-2019 : Histoire des réseaux sociaux et de leurs innovations, *Agence90*. <https://techjury.net/blog/time-spent-on-social-media/>, 11 Decembre 2019. [Accessed 04.03.2023].
- [180] Roberts Hannah. Giorgia Meloni’s Putin play, *Politico*. <https://www.politico.eu/article/giorgia-meloni-valdimir-putin-italy/>, 22 June 2022. [Accessed 18.04.2023].
- [181] Davies Harry. Ted Cruz using firm that harvested data on millions of unwitting Facebook users, *The Guardian*. <https://www.theguardian.com/us-news/2015/dec/11/senator-ted-cruz-president-campaign-facebook-user-data>, 11 September 2015. [Accessed 09.04.2023].
- [182] Bennett Holman and Justin Bruner. Experimentation by industrial selection. *Philosophy of Science*, 84(5):1008–1019, 2017.
- [183] Bennett Holman and Justin P Bruner. The problem of intransigently biased agents. *Philosophy of Science*, 82(5):956–968, 2015.
- [184] Philip N Howard, Gillian Bolsover, Bence Kollanyi, Samantha Bradshaw, and Lisa-Maria Neudert. Junk news and bots during the us election: What were michigan voters sharing

- over twitter. *CompProp, OII, Data Memo*, 1, 2017.
- [185] Philip N Howard, Bharath Ganesh, Dimitra Liotsiou, John Kelly, and Camille François. The ira, social media and political polarization in the united states, 2012-2018. 2018.
- [186] Philip N Howard, Bence Kollanyi, Samantha Bradshaw, and Lisa-Maria Neudert. Social media, news and political information during the us election: Was polarizing content concentrated in swing states? *arXiv preprint arXiv:1802.03573*, 2018.
- [187] HugoDécrypte. L’histoire derrière le succès délirant de MrBeast, *YouTube*. https://www.youtube.com/watch?v=aRUw_36wZdA, 27 November 2022. [Accessed 22.05.2023].
- [188] Consumer Insights. Attention spans, 2015.
- [189] Madruga Isabella. The Evolution Of Social Media, *The Daily Fandom*. <https://thedailyfandom.org/the-evolution-of-social-media/>, 11 April 2022. [Accessed 04.03.2023].
- [190] Labeyrie Isabelle. Ethiopie : au Tigré, la guerre est finie, mais les civils souffrent encore, *FranceTVInfo*. https://www.francetvinfo.fr/replay-radio/le-monde-est-a-nous/ethiopie-la-guerre-est-finie-mais-les-civils-souffrent-encore_5565681.html, 10 January 2023. [Accessed 23.04.2023].
- [191] Thompson Isobel. The Secret History of Steve Bannon and Alexander Nix, Explained, *Vanity Fair*. <https://www.vanityfair.com/news/2018/03/the-secret-history-of-steve-bannon-and-alexander-nix>, 21 March 2018. [Accessed 20.05.2022].
- [192] Lapowsky Issie. Facebook Doesn’t Know How Many People Followed Russians on Instagram, *Wired*. <https://www.wired.com/story/facebook-does-not-know-how-many-followers-russian-trolls-had-on-instagram/>, 02 March 2018. [Accessed 28.04.2023].
- [193] Wigmore Ivy. intermittent reinforcement, *WhatIs*. <https://www.techtarget.com/whatis/definition/intermittent-reinforcement>, May 2018. [Accessed 15.05.2022].
- [194] Goodman Jack and Carmichael Flora. Coronavirus: Bill Gates ‘microchip’ conspiracy theory and other vaccine claims fact-checked, *BBC*. <https://www.bbc.com/news/52847648>, 30 May 2020. [Accessed 28.10.2022].

-
- [195] Bulao Jacquelyn. How Much Data Is Created Every Day in 2023?, *techjury*. <https://techjury.net/blog/how-much-data-is-created-every-day/>, 07 February 2023. [Accessed 25.02.2023].
- [196] Bulao Jacquelyn. Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025, *Statista*. <https://www.statista.com/statistics/871513/worldwide-data-created/>, June 2021. [Accessed 25.02.2023].
- [197] Doward Jamie and Gibbs Alice. Did Cambridge Analytica influence the Brexit vote and the US election?, *The Guardian*. <https://www.theguardian.com/politics/2017/mar/04/nigel-oakes-cambridge-analytica-what-role-brexit-trump>, 04 March 2017. [Accessed 14.04.2023].
- [198] Mayer Jane. New Evidence Emerges of Steve Bannon and Cambridge Analytica's Role in Brexit, *The New Yorker*. <https://www.newyorker.com/news/news-desk/new-evidence-emerges-of-steve-bannon-and-cambridge-analyticas-role-in-brexit>, 17 November 2018. [Accessed 11.04.2023].
- [199] Jackson Jasper, Kassa Lucy, and Mark Townsend. Facebook 'lets vigilantes in Ethiopia incite ethnic killing', *The Guardian*. <https://www.theguardian.com/technology/2022/feb/20/facebook-lets-vigilantes-in-ethiopia-incite-ethnic-killing>, 20 February 2022. [Accessed 21.04.2023].
- [200] Jackson Jasper, Kassa Lucy, and Mark Townsend. Facebook 'lets vigilantes in Ethiopia incite ethnic killing', *The Guardian*. <https://www.theguardian.com/technology/2022/feb/20/facebook-lets-vigilantes-in-ethiopia-incite-ethnic-killing>, 20 February 2022. [Accessed 23.04.2023].
- [201] de La Gletais Jean Berthelot. Quelle est l'origine des fake news?, *Curieux*. <https://www.curieux.live/2020/12/16/quelle-est-lorigine-des-fake-news/>, 16 December 2020. [Accessed 18.02.2023].
- [202] Alric Jean-Yves. Le succès mondial des VPN se confirme (mais la France est en retard), *Press-citron*. <https://www.presse-citron.net/le-succes-mondial-des-vpn-se-confirme/>, 31 octobre 2021. [Accessed 04.10.2022].
- [203] Horwitz Jeff. Facebook Says Its Rules Apply to All. Company Documents Reveal
-

- a Secret Elite That's Exempt., *The Wall Street Journal*. https://www.wsj.com/articles/facebook-files-xcheck-zuckerberg-elite-rules-11631541353?mod=article_inline, 13 September 2021. [Accessed 21.05.2022].
- [204] Horwitz Jeff and Scheck Justin. Facebook Increasingly Suppresses Political Movements It Deems Dangerous, *The Wall Street Journal*. https://www.wsj.com/articles/facebook-suppresses-political-movements-patriot-party-11634937358?mod=article_inline, 22 October 2021. [Accessed 21.05.2022].
- [205] Gottfried Jeffrey and Shearer Elisa. News Use Across Social Media Platforms 2017, *Pew Research Center*. <https://www.pewresearch.org/journalism/2017/09/07/news-use-across-social-media-platforms-2017/>, 07 September 2017. [Accessed 18.05.2022].
- [206] Gottfried Jeffrey and Shearer Elisa. News Use Across Social Media Platforms 2016, *Pew Research Center*. <https://www.pewresearch.org/journalism/2016/05/26/news-use-across-social-media-platforms-2016/>, 26 May 2016. [Accessed 18.05.2022].
- [207] Brown Jennings. Former Facebook Exec: 'You Don't Realize It But You Are Being Programmed', *Gizmodo*. <https://gizmodo.com/former-facebook-exec-you-don-t-realize-it-but-you-are-1821181133>, 11 December 2017. [Accessed 23.05.2022].
- [208] Peters Jeremy W. Steve Bannon Steps Down From Breitbart Post, *The New York Times*. <https://www.nytimes.com/2018/01/09/us/politics/steve-bannon-breitbart-trump.html>, 09 January 2018. [Accessed 11.04.2023].
- [209] Duriez-Mise Johann. En 2004, Facebook ressemblait à ça, *Europe1*. <https://www.europe1.fr/technologies/En-2004-Facebook-ressemblait-a-ca-643360>, 04 February 2014. [Accessed 04.03.2023].
- [210] Koetsier John. 10 Most Downloaded Apps Of 2022: Facebook Down, Spotify Up, TikTok Stable, CapCut Keeps Growing, *Forbes*. <https://www.forbes.com/sites/johnkoetsier/2023/01/04/top-10-most-downloaded-apps-of-2022-facebook-down-spotify-up-tiktok-stable-capcut-> 04 January 2023. [Accessed 28.02.2023].

- [211] Koetsier John. Here Are The 10 Most Downloaded Apps Of 2020, *Forbes*. <https://www.forbes.com/sites/johnkoetsier/2021/01/07/here-are-the-10-most-downloaded-apps-of-2020/>, 07 January 2021. [Accessed 28.02.2023].
- [212] Koetsier John. Top 10 Most Downloaded Apps And Games Of 2021: TikTok, Telegram Big Winners, *Forbes*. <https://www.forbes.com/sites/johnkoetsier/2021/12/27/top-10-most-downloaded-apps-and-games-of-2021-tiktok-telegram-big-winners/>, 27 December 2021. [Accessed 28.02.2023].
- [213] Brodtkin Jon. EU names 19 large tech platforms that must follow Europe's new Internet rules, *Ars Technica*. <https://arstechnica.com/tech-policy/2023/04/google-runs-5-of-the-19-platforms-that-must-follow-eus-new-internet-rules/>, 25 April 2023. [Accessed 21.05.2022].
- [214] Swaine Jon and Bennetts Marc. Mueller charges 13 Russians with interfering in US election to help Trump, *The Guardian*. <https://www.theguardian.com/us-news/2018/feb/16/robert-mueller-russians-charged-election>, 17 February 2018. [Accessed 28.04.2023].
- [215] Michaelson Jonathan. This is Your Brain on TikTok, *Medium*. <https://medium.com/@jonathanmichaelson/this-is-your-brain-on-tiktok-8415583446eb>, 13 November 2022. [Accessed 23.05.2022].
- [216] Novet Jordan. Microsoft is laying off 10,000 employees, *CNBC*. <https://www.cnbc.com/2023/01/18/microsoft-is-laying-off-10000-employees.html>, 18 January 2023. [Accessed 17.05.2022].
- [217] T. Prodanoff Jordan. How Much Data Is Created Every Day in 2022, *Web tribunal*. <https://webtribunal.net/blog/how-much-data-is-created-every-day/>, 07 October 2022. [Accessed 24.02.2023].
- [218] Schiller Judah. Average Human Attention Span By Age: 31 Statistics, *The treetop therapy*. <https://www.thetreetop.com/statistics/average-human-attention-span>, 30 March 2023. [Accessed 06.04.2023].
- [219] Carrie Wong Julia. How Facebook let fake engagement distort global politics: a whistleblower's account, *The Guardian*. <https://www.theguardian.com/technology/2021/>

- apr/12/facebook-fake-engagement-whistleblower-sophie-zhang, 12 April 2021. [Accessed 11.12.2022].
- [220] Carrie Wong Julia. Former Facebook executive: social media is ripping society apart, *The Guardian*. <https://www.theguardian.com/technology/2017/dec/11/facebook-former-executive-ripping-society-apart>, 12 December 2017. [Accessed 03.04.2023].
- [221] Wong Julia Carrie and Ernst Jeff. Facebook knew of Honduran president's manipulation campaign and let it continue for 11 months, *The Guardian*. <https://www.theguardian.com/technology/2021/apr/13/facebook-honduras-juan-orlando-hernandez-fake-engagement>, 13 April 2021. [Accessed 16.05.2022].
- [222] Lacaze Julie. L'histoire controversée des Rohingyas en Birmanie, *National Geographic*. <https://www.nationalgeographic.fr/histoire/lhistoire-controversee-des-rohingyas-en-birmanie>, No Date. [Accessed 21.04.2023].
- [223] Ling Justin. The Fake News About Cambridge Analytica, *The Walrus*. <https://thewalrus.ca/the-fake-news-about-cambridge-analytica/>, 05 February 2020. [Accessed 20.05.2022].
- [224] Scheck Justin, Purnell Newley, and Jeff Horwitz. Facebook Employees Flag Drug Cartels and Human Traffickers. The Company's Response Is Weak, Documents Show., *The Wall Street Journal*. https://www.wsj.com/articles/facebook-drug-cartels-human-traffickers-response-is-weak-documents-11631812953?mod=article_inline, 16 September 2021. [Accessed 21.04.2023].
- [225] justLikeAPI. How does Facebook detect bots?, *Medium*. <https://justlikeapi.medium.com/how-does-facebook-detect-bots-ea2481ac24de>, 23 Jun 2023. [Accessed 24.05.2023].
- [226] Hill Kashmir. How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did, *Forbes*. <https://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>, 16 February 2012. [Accessed 18.03.2023].

- [227] Hill Kashmir. Facebook Manipulated 689,003 Users' Emotions For Science, *Forbes*. <https://www.forbes.com/sites/kashmirhill/2014/06/28/facebook-manipulated-689003-users-emotions-for-science/>, 28 Jun 2014. [Accessed 16.03.2023].
- [228] Eva Mutsaers Katerina and Shearer Elisa. News Use Across Social Media Platforms 2018, *Pew Research Center*. <https://www.pewresearch.org/journalism/2018/09/10/news-use-across-social-media-platforms-2018/>, 10 September 2018. [Accessed 18.05.2022].
- [229] McQue Katie and McNamara Mei-Ling. How Facebook and Instagram became marketplaces for child sex trafficking, *The Guardian*. <https://www.theguardian.com/news/2023/apr/27/how-facebook-and-instagram-became-marketplaces-for-child-sex-trafficking>, 27 April 2023. [Accessed 23.05.2022].
- [230] Hagey Keach and Horwitz Jeff. Facebook's Internal Chat Boards Show Politics Often at Center of Decision Making, *The Wall Street Journal*. https://www.wsj.com/articles/facebook-politics-decision-making-documents-11635100195?mod=article_inline, 24 October 2021. [Accessed 21.05.2022].
- [231] Dilanian Ken. Russian trolls who interfered in 2016 U.S. election also made ad money, report says, *NBC News*. <https://www.nbcnews.com/politics/national-security/russian-trolls-who-interfered-2016-u-s-election-also-made-n1013811>, 05 June 2019. [Accessed 28.04.2023].
- [232] Vogel Kenneth P. and Parti Tarini. Cruz partners with donor's 'psychographic' firm, *Politico*. <https://www.politico.com/story/2015/07/ted-cruz-donor-for-data-119813>, 07 Juny 2015. [Accessed 14.04.2023].
- [233] Deniau K vin. Cambridge Analytica : tout comprendre sur la plus grande crise de l'histoire de Facebook, *SiecleDigital*. <https://siecledigital.fr/2018/03/23/cambridge-analytica-tout-comprendre-sur-la-plus-grande-crise-de-lhistoire-de-facebook>, 23 March 2018. [Accessed 07.04.2023].
- [234] McSpadden Kevin. You Now Have a Shorter Attention Span Than a Goldfish, *Time*. <https://time.com/3858309/attention-spans-goldfish/>, 14 May 2015. [Accessed 06.04.2023].

- [235] Roose Kevin. Russian Trolls Came for Instagram, Too, *The New York Times*. <https://www.nytimes.com/2018/12/18/technology/russian-interference-instagram.html>, 18 December 2018. [Accessed 28.04.2023].
- [236] Adam DI Kramer, Jamie E Guillory, and Jeffrey T Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24):8788–8790, 2014.
- [237] Gambino Lauren. Facebook says up to 10m people saw ads bought by Russian agency, *The Guardian*. <https://www.theguardian.com/technology/2017/oct/02/facebook-says-up-to-10m-people-saw-ads-bought-by-russian-agency>, 03 October 2017. [Accessed 28.04.2023].
- [238] David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, et al. The science of fake news. *Science*, 359(6380):1094–1096, 2018.
- [239] Victor Le Pochat, Laura Edelson, Tom Van Goethem, Wouter Joosen, Damon McCoy, and Tobias Lauinger. An audit of facebook’s political ad policy enforcement. In *31st USENIX Security Symposium (USENIX Security 22)*, pages 607–624, 2022.
- [240] L  a. Quelle est la part du trafic mobile sur Internet en 2022 ?, *Statista*. <https://www.statista.com/statistics/272697/market-share-desktop-internet-browser-usa/>, 16 February 2022. [Accessed 28.02.2023].
- [241] Hogan in Yangon Libby and Michael Safi. Revealed: Facebook hate speech exploded in Myanmar during Rohingya crisis, *The Guardian*. <https://www.theguardian.com/world/2018/apr/03/revealed-facebook-hate-speech-exploded-in-myanmar-during-rohingya-crisis>, 03 April 2018. [Accessed 21.04.2023].
- [242] Rondaut Lucie. Quelle est la difference entre le Web et Internet, *Le Figaro*. <https://www.lefigaro.fr/secteur/high-tech/2019/03/11/32001-20190311ARTFIG00004-quelle-est-la-difference-entre-le-web-et-internet.php>, 11 March 2019. [Accessed 24.02.2023].
- [243] Turcotte Marc-Antoine. Confusion entre le coronavirus et la bi  re Corona,

- Le Journal de Montréal*. <https://www.journaldemontreal.com/2020/01/29/confusion-entre-le-coronavirus-et-la-biere-corona>, 29 January 2020. [Accessed 16.10.2022].
- [244] Zuckerberg Mark. Mark Zuckerberg’s Message to Meta Employees, *Meta*. <https://about.fb.com/news/2022/11/mark-zuckerberg-layoff-message-to-employees/>, 09 November 2022. [Accessed 17.05.2022].
- [245] Chan Marsha. The Psychology of Social Media, *Medium*. <https://medium.com/@MarshaChan/the-psychology-of-social-media-712222f07d1d>, 27 February 2016. [Accessed 05.04.2023].
- [246] Untersinger Martin and Leloup Damien. Facebook emploie 766 modérateurs en langue arabe pour 220 millions d’utilisateurs arabophones, *Le Monde*. https://www.lemonde.fr/pixels/article/2021/11/16/facebook-emploie-766-moderateurs-en-langue-arabe-pour-220-millions-d-utilisateurs-6102312_4408996.html, 16 November 2021. [Accessed 22.04.2023].
- [247] Zhang Mary. Facebook’s 18 Data Centers: \$20bn Investment, 40m Square Feet, *Dgtl Infra*. <https://dgtlinfra.com/facebook-18-data-centers-20bn-investment/>, 27 September 2017. [Accessed 24.02.2023].
- [248] Komnenic Masha. Cookie Law Guide for Businesses: EU, US, and the UK, *Termly*. <https://termly.io/resources/articles/cookie-law/>, 21 February 2018. [Accessed 05.10.2022].
- [249] Walker Mason and Eva Matsa Katerina. News Consumption Across Social Media in 2021, *Pew Research Center*. <https://www.pewresearch.org/journalism/2021/09/20/news-consumption-across-social-media-in-2021/>, 20 September 2021. [Accessed 18.05.2022].
- [250] Burgess Matt. We finally know the full extent of Russia’s Twitter trolling campaign, *Wired*. <https://www.wired.co.uk/article/twitter-troll-data-russia-ira-iran>, 17 October 2018. [Accessed 28.04.2023].
- [251] Holroyd Matthew. Italy elections: Who has spent the most money on Facebook ads?, *Euronews*. <https://www.euronews.com/my-europe/2022/09/23/italy-elections-who-has-spent-the-most-money-on-facebook-ads>, 23 Septem-

- ber 2022. [Accessed 18.04.2023].
- [252] Solan Matthew. The secret to happiness? Here’s some advice from the longest-running study on happiness, *Harvard Health Publishing*. <https://www.health.harvard.edu/blog/the-secret-to-happiness-heres-some-advice-from-the-longest-running-study-on-happiness>, 05 October 2017. [Accessed 05.04.2023].
- [253] Roser Max. The Internet’s history has just begun, *Our World in Data*. <https://ourworldindata.org/internet-history-just-begun>, 03 October 2018. [Accessed 24.02.2023].
- [254] Dar Meshi, Carmen Morawetz, and Hauke R Heekeren. Nucleus accumbens response to gains in reputation for the self relative to gains for others predicts social media use. *Frontiers in human neuroscience*, page 439, 2013.
- [255] Dimock Michael, Kiley Jocelyn, Keeter Scott, and Carroll Doherty. Political Polarization in the American Public, *Pew Research Center*. <https://www.pewresearch.org/politics/2014/06/12/political-polarization-in-the-american-public/>, 12 June 2014. [Accessed 09.11.2022].
- [256] Krever Mick. Wagner chief admits to founding Russian troll farm sanctioned for meddling in US elections, *CNN*. <https://edition.cnn.com/2023/02/14/europe/russia-yevgeny-prigozhin-internet-research-agency-intl/index.html>, 14 February 2023. [Accessed 28.04.2023].
- [257] Allen Mike. Sean Parker unloads on Facebook: “God only knows what it’s doing to our children’s brains”, *Axios*. <https://www.axios.com/2017/12/15/sean-parker-unloads-on-facebook-god-only-knows-what-its-doing-to-our-childrens-brains>, 09 November 2017. [Accessed 19.03.2023].
- [258] Isaac Mike. Meta to Lay Off Another 10,000 Workers, *The New York Times*. <https://www.nytimes.com/2023/03/14/technology/meta-facebook-layoffs.html>, 14 March 2023. [Accessed 17.05.2022].
- [259] Isaac Mike and Wakabayashi Daisuke. Russian Influence Reached 126 Million Through Facebook Alone, *The New York Times*. <https://www.nytimes.com/2017/10/30/technology/facebook-google-russia.html>, 30 October 2017. [Accessed 28.04.2023].

- [260] Mina. Quelques statistiques sur les VPN, *VPN mon ami*. <https://www.vpnmonami.com/quelques-statistiques-sur-les-vpn/>, 07 march 2022. [Accessed 04.10.2022].
- [261] Monica Murero. Coordinated inauthentic behavior: An innovative manipulation tactic to amplify covid-19 anti-vaccine communication outreach via social media. *Frontiers in Sociology*, 8, 2023.
- [262] Gleicher Nathaniel. Coordinated Inauthentic Behavior Explained, *Meta*. <https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/>, 06 December 2018. [Accessed 09.05.2023].
- [263] MacFarquhar Neil. Inside the Russian Troll Factory: Zombies and a Breakneck Pace, *The New York Times*. <https://www.nytimes.com/2018/02/18/world/europe/russia-troll-factory.html>, 18 February 2018. [Accessed 24.04.2023].
- [264] Netflix. The social dilemma. <https://www.netflix.com/be-fr/title/81254224>, 9 September 2020.
- [265] Purnell Newley and Horwitz Jeff. Facebook Services Are Used to Spread Religious Hatred in India, Internal Documents Show, *The Wall Street Journal*. https://www.wsj.com/articles/facebook-services-are-used-to-spread-religious-hatred-in-india-internal-documents-mod=article_inline, 23 October 2021. [Accessed 21.05.2022].
- [266] Breitbart News. Blow-Out: Breitbart News Sets Traffic Records in July 192 Million Pageviews, 31 Million Uniques, 89 Million Visits, *Breitbart*. <https://www.breitbart.com/the-media/2016/08/07/blow-breitbart-news-sets-traffic-records-july-192-million-pageviews-31-million-uni>, 06 August 2016. [Accessed 11.04.2023].
- [267] Confessore Nicholas. Cambridge Analytica and Facebook: The Scandal and the Fallout So Far, *The New York Time*. <https://www.nytimes.com/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html>, 04 April 2018. [Accessed 14.04.2023].
- [268] Confessore Nicholas. How One Family’s Deep Pockets Helped Reshape Donald Trump’s Campaign, *The New York Times*. <https://www.nytimes.com/2016/08/19/us/politics/robert-mercere-donald-trump-donor.html>, 18 August 2018. [Accessed 11.04.2023].

- [269] Thompson Nicholas and Fred Vogelstein. Inside the Two Years That Shook Facebook and the World, *Wired*. <https://www.wired.com/story/inside-facebook-mark-zuckerberg-2-years-of-hell/>, 12 February 2018. [Accessed 11.04.2023].
- [270] Bilton Nick. Steve Job was a low-tech parent, *The New York Times*. <https://www.nytimes.com/2014/09/11/fashion/steve-jobs-apple-was-a-low-tech-parent.html>, 10 September 2014. [Accessed 23.05.2022].
- [271] Lellouche Nicolas. Mark Zuckerberg va licencier 10 000 employés Meta, après en avoir viré 11 000, *Numerama*. <https://www.numerama.com/tech/1303198-mark-zuckerberg-va-licencier-10-000-employees-meta-apres-en-avoir-vire-11-000.html>, 14 March 2023. [Accessed 17.05.2022].
- [272] Six Nicolas. Massacre des Rohingya : Facebook a joué un rôle central dans la montée du climat de haine en Birmanie, *Le Monde*. https://www.lemonde.fr/pixels/article/2022/09/29/massacre-des-rohingya-facebook-a-joue-un-role-central-dans-la-montee-du-climat-de-birmanie_6143611_4408996.html, 29 September 2022. [Accessed 21.04.2023].
- [273] Christopher Niles. Facebook’s language problem, *Rest Of World*. <https://restofworld.org/2021/newsletter-south-asia-facebooks-language-problem/>, 02 November 2021. [Accessed 23.04.2023].
- [274] Maudet Nolwenn. Eli Pariser, The filter bubble, Penguin Books, 2011, *Unilim*. <https://www.unilim.fr/interfaces-numeriques/1775>, 12 January 2018. [Accessed 24.05.2022].
- [275] Cyrus North. La surenchère des surproductions YouTube, *YouTube*. <https://www.youtube.com/watch?v=XgjyRVwHQQc>, 21 May 2023. [Accessed 22.05.2023].
- [276] Cailin O’Connor and Justin Bruner. Dynamics and diversity in epistemic communities. *Erkenntnis*, 84:101–119, 2019.
- [277] The Editors of Encyclopaedia Britannica. Robert Mueller - American law enforcement official, *Britannica*. <https://www.britannica.com/biography/Robert-Mueller>, 16 May 2023. [Accessed 24.05.2022].
- [278] Djuraskovic Ogi. Big Data Statistics 2023: How Much Data is in The World?, *First*

- Site Guide*. <https://firstsiteguide.com/big-data-stats/>, 26 December 2022. [Accessed 25.02.2023].
- [279] Solon Olivia. Ex-Facebook president Sean Parker: site made to exploit human 'vulnerability', *The Guardian*. <https://www.theguardian.com/technology/2017/nov/09/facebook-sean-parker-vulnerability-brain-psychology>, 09 November 2017. [Accessed 19.03.2023].
- [280] Solon Olivia. Twitter plans to make political ads more transparent amid Russia revelations, *The Guardian*. <https://www.theguardian.com/technology/2017/oct/24/twitter-transparency-political-ads-russia-us-election-trump>, 24 October 2017. [Accessed 28.04.2023].
- [281] Solon Olivia and Siddiqui Sabrina. Russia-backed Facebook posts 'reached 126m Americans' during US election, *The Guardian*. <https://www.theguardian.com/technology/2017/oct/30/facebook-russia-fake-accounts-126-million>, 31 October 2017. [Accessed 28.04.2023].
- [282] Esteban Ortiz-Ospina. The rise of social media. *Our World in Data*, 2019. <https://ourworldindata.org/rise-of-social-media>.
- [283] Pinnell Owen and Kelly Jess. Slave markets found on Instagram and other apps, *BBC*. <https://www.bbc.com/news/technology-50228549>, 31 October 2019. [Accessed 17.05.2022].
- [284] Martineau Paris. How Instagram Became the Russian IRA's Go-To Social Network, *Wired*. <https://www.wired.com/story/how-instagram-became-russian-iras-social-network/>, 17 December 2018. [Accessed 28.04.2023].
- [285] Khal Patrice. Pourquoi Facebook a-t-il racheté Whatsapp ?, *Swello*. <https://swello.com/fr/blog/rachat-whatsapp/>, 03 January 2019. [Accessed 04.03.2023].
- [286] Kirby Paul. Giorgia Meloni: Italy's far-right wins election and vows to govern for all, *BBC*. <https://www.bbc.com/news/world-europe-63029909>, 26 September 2022. [Accessed 18.04.2023].
- [287] Lewis Paul. 'Our minds can be hijacked': the tech insiders who fear a smart-phone dystopia, *The Guardian*. <https://www.theguardian.com/technology/2017/>

- oct/05/smartphone-addiction-silicon-valley-dystopia, 06 October 2017. [Accessed 04.04.2023].
- [288] Dizikes Peter. Study: On Twitter, false news travels faster than true stories, *MIT News Office*. <https://news.mit.edu/2018/study-twitter-false-news-travels-faster-true-stories-0308>, 8 March 2018. [Accessed 16.10.2022].
- [289] Hraka Peter. We've analysed 500,000 browser fingerprints. Here is what we found., *Medium*. <https://medium.com/slido-dev-blog/we-collected-500-000-browser-fingerprints-here-is-what-we-found-82c319464dc9>, 7 February 2019. [Accessed 05.10.2022].
- [290] Alonso Pierre. Cambridge Analytica, la pieuvre de l'extrême droite américaine, *Liberation*. https://www.liberation.fr/planete/2018/04/10/cambridge-analytica-la-pieuvre-de-l-extreme-droite-americaine_1642499/, 10 April 2018. [Accessed 20.05.2022].
- [291] Danse Pierre-François. Qu'est-ce qu'un growth-hacker ? Définition & exemples, *Strat-net*. <https://blog.stratenet.com/inspiration/quest-ce-quun-growth-hacker>, 02 Septembre 2014. [Accessed 12.10.2022].
- [292] Franck Rebillard. La rumeur du pizzagate durant la présidentielle de 2016 aux états-unis. *Réseaux*, 202203(2):273–310, 2017.
- [293] Staff Reuters. Big moments in Mueller investigation of Russian meddling in 2016 U.S. election, *Reuters*. <https://www.reuters.com/article/uk-usa-trump-russia-timeline-idUKKCN1QF2DA>, 26 February 2019. [Accessed 28.04.2023].
- [294] Gonzales Richard. Steve Bannon Out At Breitbart News, *The Two-Way*. <https://www.npr.org/sections/thetwo-way/2018/01/09/576868468/bannon-out-at-breitbart-news>, 09 Januray 2018. [Accessed 11.04.2023].
- [295] Barry Rob, Wells Georgia, West John, and Stern Joanna. How TikTok Serves Up Sex and Drug Videos to Minors, *The Wall Street Journal*. <https://www.wsj.com/articles/tiktok-algorithm-sex-drugs-minors-11631052944>, 08 September 2021. [Accessed 21.05.2022].

-
- [296] Booth Robert. Facebook reveals news feed experiment to control emotions, *The Guardian*. <https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds>, 30 Jun 2014. [Accessed 16.03.2023].
- [297] Mueller Robert. IN THE UNITED STATES DISTRICT COURT FOR THE DISTRICT OF COLUMBIA, *The Guardian*. <https://www.justice.gov/archives/sco/file/1035477/download>, 16 February 2018. [Accessed 28.04.2023].
- [298] Helderman Rosalind S. Here's what we know about Donald Trump and his ties to Russia, *The Washington Post*. <https://www.axios.com/2019/04/18/mueller-report-trump-russia-interference-collusion>, 29 July 2016. [Accessed 16.05.2022].
- [299] Sini Rozina. 'You are being programmed,' former Facebook executive warns, *BBC News*. <https://www.bbc.com/news/blogs-trending-42322746>, 12 December 2017. [Accessed 03.04.2023].
- [300] Hannah Rubin, Justin P Bruner, Cailin O'Connor, and Simon Huttegger. Communication without common interest: A signaling experiment. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 83:101295, 2020.
- [301] Derek Ruths and Jürgen Pfeffer. Social media for large studies of behavior. *Science*, 346(6213):1063–1064, 2014.
- [302] broderick Ryan. Here's Everything The Mueller Report Says About How Russian Trolls Used Social Media, *BuzzFeed News*. <https://www.buzzfeednews.com/article/ryanhatesthis/mueller-report-internet-research-agency-detailed-2016>, 18 April 2019. [Accessed 28.04.2023].
- [303] Mac Ryan. Instagram Censored Posts About One Of Islam's Holiest Mosques, Drawing Employee Ire, *BuzzFeed News*. <https://www.buzzfeednews.com/article/ryanmac/instagram-facebook-censored-al-aqsa-mosque>, 12 May 2021. [Accessed 22.04.2023].
- [304] Dixon S. Annual advertising revenue of Meta Platforms worldwide from 2009 to 2022, *Statista*. <https://www.statista.com/statistics/271258/facebooks-advertising-revenue-worldwide/>, 13 February 2023. [Accessed

07.04.2023].

- [305] Dixon S. Leading countries based on Facebook audience size as of January 2023, *Statista*. <https://www.statista.com/statistics/268136/top-15-countries-based-on-number-of-facebook-users/>, 24 February 2023. [Accessed 21.05.2022].
- [306] Knight Sam. Life Inside S.C.L., Cambridge Analytica’s Parent Company, *The New Yorker*. <https://www.newyorker.com/news/letter-from-the-uk/life-inside-scl-cambridge-analyticas-parent-company>, 26 March 2018. [Accessed 14.04.2023].
- [307] Levin Sam. Facebook told advertisers it can identify teens feeling ‘insecure’ and ‘worthless’, *The Guardian*. <https://www.theguardian.com/technology/2017/may/01/facebook-advertising-data-insecure-teens>, 01 May 2017. [Accessed 04.04.2023].
- [308] Gibbs Samuel. Apple’s Tim Cook: ‘I don’t want my nephew on a social network’, *The Guardian*. <https://www.theguardian.com/technology/2018/jan/19/tim-cook-i-dont-want-my-nephew-on-a-social-network>, 19 Janvier 2018. [Accessed 14.10.2022].
- [309] Laurent Samuel. Il y a 40 ans naissait (presque) Internet, *Le Figaro*. <https://www.lefigaro.fr/web/2009/09/02/01022-20090902ARTFIG00263-il-y-a-40-ans-naissait-presque-internet-.php>, 02 September 2009. [Accessed 24.02.2023].
- [310] Faes Parisi Sandro. 11.000 employ  s vir  s. Meta, la maison m  re de Facebook, Instagram et WhatsApp, licencie 13% de ses effectifs, *RTBF*. <https://www.rtbef.be/article/11000-employees-vires-meta-la-maison-mere-de-facebook-instagram-et-whatsapp-licenci>, 09 November 2022. [Accessed 17.05.2022].
- [311] Salinas Sara. Facebook co-founder Sean Parker bashes company, saying it was built to exploit human vulnerability, *CNBC*. <https://www.cnn.com/2017/11/09/facebooks-sean-parker-on-social-media.html>, 09 November 2017. [Accessed 19.03.2023].
- [312] Shane Scott and Frenkel Sheera. Russian 2016 Influence Operation Targeted African-Americans on Social Media, *The New York Time*. <https://www.nytimes.com/2018/>

- 12/17/us/politics/russia-2016-influence-campaign.html, 17 Decembre 2018. [Accessed 28.04.2023].
- [313] Shane Scott and Goel Vindu. Fake Russian Facebook Accounts Bought \$100,000 in Political Ads, *The New York Times*. <https://www.nytimes.com/2017/09/06/technology/facebook-russian-political-ads.html>, 06 September 2017. [Accessed 28.04.2023].
- [314] Arabi Shahida. Intermittent Reinforcement: The Powerful Manipulation Method That Keeps You Trauma Bonded To Your Abuser, *Thought Catalog*. <https://thoughtcatalog.com/shahida-arabi/2017/11/this-powerful-manipulation-method-keeps-you-bonded-to-your-abuser/>, 1 February 2022. [Accessed 12.10.2022].
- [315] Mamou Sha  . Kyrie Irving s  excuse pour l  affaire de la Terre plate, *BasketSession*. <https://www.basketsession.com/NBA/kyrie-irving-sexcuse-pour-laffaire-de-la-terre-plate-437269/>, 02 October 2018. [Accessed 15.10.2022].
- [316] Walker Shaun and Sauer Pjotr. Yevgeny Prigozhin: the hot-dog seller who rose to the top of Putin  s war machine, *The Guardian*. <https://www.theguardian.com/world/2023/jan/24/yevgeny-prigozhin-the-hotdog-seller-who-rose-to-the-top-of-putin-war-machine-wagne>, 24 January 2023. [Accessed 09.05.2023].
- [317] Frenkel Sheera and Alba Davey. In India, Facebook Grapples With an Amplified Version of Its Problems, *The New York Times*. <https://www.nytimes.com/2021/10/23/technology/facebook-india-misinformation.html>, 23 October 2021. [Accessed 23.04.2023].
- [318] Frenkel Sheera, Rosenberg Matthew, and Confessore Nicholas. How Trump Consultants Exploited the Facebook Data of Millions, *The New York Times*. <https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html?mtrref=www.lesechos.fr>, 17 March 2017. [Accessed 07.04.2023].
- [319] Matthew Rosenberg Sheera Frenkel and Nicholas Confessore. Facebook data collected by quiz app included private messages, *The New York Times*. <https://www.nytimes.com/2018/04/10/technology/facebook-cambridge-analytica-private-messages.html>, 10 April 2018. [Accessed 07.04.2023].

- [320] Goel Shikhar. How does Facebook (Meta) make money: Business Model Analysis, *The Strategy Story*. https://thestrategystory.com/2023/01/09/how-does-facebook-meta-make-money-business-model-analysis/?utm_source=rss&utm_medium=rss&utm_campaign=how-does-facebook-meta-make-money-business-model-a 09 January 2023. [Accessed 07.04.2023].
- [321] Zuboff Shoshana. Votre brosse à dents vous espionne. Un capitalisme de surveillance, *Le Monde diplomatique*. <https://www.monde-diplomatique.fr/2019/01/ZUBOFF/59443>, January 2019. [Accessed 15.03.2023].
- [322] Zuboff Shoshana. They have no right to my face or yours., *Shoshana Zuboff*. <https://shoshanazuboff.com/book/shoshana/>, No date. [Accessed 29.05.2023].
- [323] Kemp Simon. Digital 2017: Honduras, *DATAREPORTAL*. <https://datareportal.com/reports/digital-2017-honduras>, 01 February 2017. [Accessed 16.05.2022].
- [324] Kemp Simon. Digital 2022: April Global Statshot Report, *DataReportal*. <https://datareportal.com/reports/digital-2022-april-global-statshot>, 24 April 2022. [Accessed 24.02.2023].
- [325] Amsili Sophie. Comprendre la crise des Rohingyas en quatre questions, *Les Echos*. <https://www.lesechos.fr/monde/asie-pacifique/comprendre-la-crise-des-rohingyas-en-quatre-questions-1165414>, 23 January 2020. [Accessed 21.04.2023].
- [326] Bécherel Sophie. 7 000 langues sont parlées dans le monde mais 25 disparaissent chaque année, *Radiofrance*. <https://www.radiofrance.fr/franceinter/7-000-langues-sont-parlees-dans-le-monde-mais-25-disparaissent-chaque-annee-579537> 07 February 2021. [Accessed 22.04.2023].
- [327] Staff. A New Dawn in Digital: Neuromarketing Meets Social Media, *WeRSM*. <https://wersm.com/a-new-dawn-in-digital-neuromarketing-meets-social-media/>, 11 February 2015. [Accessed 20.05.2023].
- [328] Staff and agencies. 'Troll factory' spreading Russian pro-war lies online, says UK, *The Guardian*. <https://www.theguardian.com/world/2022/may/01/troll-factory-spreading-russian-pro-war-lies-online-says-uk>, 01 May 2022. [Accessed 28.04.2023].

- [329] Ozsvath Stephan. Hungary's Fidesz party seeks to conquer social media, *DW*. <https://www.dw.com/en/hungarys-fidesz-party-seeks-to-conquer-social-media/a-56479198>, 02 June 2022. [Accessed 18.04.2023].
- [330] Gadais Tegwen. Voici pourquoi nous avons tant besoin les uns des autres, *The Conversation*. <https://theconversation.com/voici-pourquoi-nous-avons-tant-besoin-les-uns-des-autres-139729>, 04 June 2020. [Accessed 05.04.2023].
- [331] Gross Terry. Whistleblower Explains How Cambridge Analytica Helped Fuel U.S. 'Insurgency', *npr*. <https://www.npr.org/2019/10/08/768216311/whistleblower-explains-how-cambridge-analytica-helped-fuel-u-s-insurgency>, 08 October 2019. [Accessed 24.05.2022].
- [332] A. Th. Le coronavirus porte préjudice à la bière Corona, *Le Soir*. <https://www.lesoir.be/283174/article/2020-02-27/le-coronavirus-porte-prejudice-la-biere-corona>, 27 February 2020. [Accessed 16.10.2022].
- [333] Trauchessec Theo. Tout savoir sur les Reels Instagram, *YourComics*. <https://your-comics.com/fr/article/definition-reel-instagram-62.html>, 27 January 2022. [Accessed 15.05.2022].
- [334] Roudet Thomas. Comprendre le Growth Hacking en 5 minutes, *La fabrique du net*. <https://www.lafabriquedunet.fr/blog/comprendre-growth-hacking-5-minutes/>, 21 march 2016. [Accessed 12.10.2022].
- [335] Bianchi Tiago. Advertising revenue of Google from 2001 to 2022, *Statista*. <https://www.statista.com/statistics/266249/advertising-revenue-of-google/>, 24 February 2023. [Accessed 07.04.2023].
- [336] Gerken Tom. Meta lay-offs: Facebook owner to cut 10,000 staff, *BBC*. <https://www.bbc.com/news/technology-64954124>, 14 March 2023. [Accessed 17.05.2022].
- [337] Romm Tony. 'Pro-Beyoncé' vs. 'Anti-Beyoncé': 3,500 Facebook ads show the scale of Russian manipulation, *The Washington Post*. <https://www.washingtonpost.com/news/the-switch/wp/2018/05/10/here-are-the-3400-facebook-ads-purchased-by-russias-online-trolls-during-the-2016->

- 10 May 2018. [Accessed 28.04.2023].
- [338] Gaudiaut Tristan. Comment les géants de la tech gagnent-ils leurs milliards ?, *Statista*. <https://fr.statista.com/infographie/13194/principales-sources-de-revenus-gafam-geants-tech-part-segment-activites-produits-d> 02 May 2022. [Accessed 07.04.2023].
- [339] Gaudiaut Tristan. Les géants de la publicité en ligne, *Statista*. <https://fr.statista.com/infographie/28966/principales-societes-plateformes-selon-les-ventes-mondiales-de-publicite-en-ligne-> 15 December 2022. [Accessed 07.04.2023].
- [340] Gaudiaut Tristan. YouTube génère plus de 20% du trafic Internet mobile, *Statista*. <https://fr.statista.com/infographie/17351/applications-generant-la-plus-grande-part-du-traffic-internet-mobile/>, 27 September 2021. [Accessed 28.02.2023].
- [341] Shukla Utkarsh. Attention Span: Meaning And Significance, *Tutorials point*. <https://www.tutorialspoint.com/attention-span-meaning-and-significance>, 09 December 2022. [Accessed 06.04.2023].
- [342] Joly Vinciane. Section 230 : la Cour suprême américaine examine le texte clé de la liberté d’expression sur le Web, *La Croix*. <https://www.la-croix.com/Monde/Section-230-Cour-supreme-americaine-examine-texte-cle-liberte-dexpression-Web-2023> 21 February 2023. [Accessed 23.05.2022].
- [343] Shevchenko Vitaly. Yevgeny Prigozhin: From Putin’s chef to Wagner founder, *BBC News*. <https://www.bbc.com/news/world-europe-64976080>, 20 March 2023. [Accessed 09.05.2023].
- [344] Elliott Vittoria, Christopher Nilesh, Deck Andrew, and Schwartz Leo. The Facebook Papers reveal staggering failures in the Global South, *Rest Of World*. <https://restofworld.org/2021/facebook-papers-reveal-staggering-failures-in-global-south/>, 26 October 2021. [Accessed 23.04.2023].
- [345] Makszimov Vlad. Orban’s influencers shower cash, become largest social media spenders, *Euractiv*. https://www.euractiv.com/section/politics/short_news/

- orbans-influencers-shower-cash-become-largest-social-media-spenders/, 04 January 2022. [Accessed 18.04.2023].
- [346] Soroush Vosoughi, Deb Roy, and Sinan Aral. The spread of true and false news online. *science*, 359(6380):1146–1151.
- [347] Adrian F Ward, Kristen Duke, Ayelet Gneezy, and Maarten W Bos. Brain drain: The mere presence of one’s own smartphone reduces available cognitive capacity. *Journal of the Association for Consumer Research*, 2(2):140–154, 2017.
- [348] James Owen Weatherall, Cailin O’Connor, and Justin P Bruner. How to beat science and influence people: policymakers and propaganda in epistemic networks. *The British Journal for the Philosophy of Science*, 2020.
- [349] Jen Weedon, William Nuland, and Alex Stamos. Information operations and facebook. Retrieved from Facebook: <https://fbnewsroomus.files.wordpress.com/2017/04/facebook-and-information-operations-v1.pdf>, 2017.
- [350] Siegelman Wendy. SCL Group - Companies & Shareholders, *Medium*. <https://wsiegelman.medium.com/scl-companies-shareholders-e65a4f394158>, 09 May 2017. [Accessed 09.05.2023].
- [351] Grandmontagne Yves. L’attention de l’humain est devenue inférieure à celle du poisson rouge!, *It Social*. <https://itsocial.fr/enjeux-it/enjeux-metiers/marketing/lattention-de-lhumain-devenue-inferieure-a-celle-poisson-rouge>, 14 September 2016. [Accessed 06.04.2023].
- [352] ROBIN Yves-Marie. France Télévisions. Il n’y aura plus de pubs dans les émissions pour les enfants, *Ouest France*. <https://www.ouest-france.fr/sante/france-televisions-il-n-y-aura-plus-de-pubs-dans-les-emissions-pour-les-enfants-54>, 29 December 2017. [Accessed 21.05.2022].
- [353] Basu Zachary. What the Mueller report tells us about Trump and Russia, *Axios*. <https://www.axios.com/2019/04/18/mueller-report-trump-russia-interference-collusion>, 18 April 2019. [Accessed 16.05.2022].
- [354] Kevin JS Zollman. The communication structure of epistemic communities. *Philosophy of science*, 74(5):574–587, 2007.

-
- [355] Szelényi Zsuzsanna. La relation privilégiée entre Meloni et Orbn va-t-elle changer la politique européenne?, *Libération*. <https://www.liberation.fr/idees-et-debats/tribunes/la-relation-privilegiee-entre-meloni-et-orban-va-t-elle-changer-la-politique-europ/JZPTXADVARBKDNVA44WPF6ZTJE/>, 04 November 2022. [Accessed 18.04.2023].
- [356] Shoshana Zuboff. Big other: surveillance capitalism and the prospects of an information civilization. *Journal of information technology*, 30(1):75–89, 2015.
- [357] Shoshana Zuboff and Jonathan Chalker. Le capitalisme de la surveillance: Un nouveau clergé. *Esprit*, (5):63–77, 2019.

A. Source code of the experiments

We have included the codes of the various experiments in case you would like to continue them.

A.1. Codes for the first experiment

This section consists of four files: Module.py, Bot_test.py, NewBot.py and Script_Navigation_bot.py.

1. The `Module.py` file contains various modules used in the program. This file is imported into other files.
2. The file `Bot_test.py` serves as a testing file that helped us familiarize ourselves with the Selenium module and understand its usage. Once we gained a solid understanding, we proceeded to create a bot class and another file that utilizes this class. These two additional files will be presented next.
3. The file `NewBot.py` contains our Bot class, which incorporates various methods for a bot.
4. The `Script_Navigation_bot.py` file utilizes our NewBot class, establishes a connection to Facebook, and performs scrolling to gather various information from 5 successive posts.

It is important to mention that the files are not fully completed. As previously explained in chapter 2, we came to the realization that this approach was not viable. Therefore, we halted the coding process and redirected our focus towards another method. As a result, there are sections in the code that are marked as *TO DO*. Additionally, certain parts of the code were still in the development stage. So, there are sections that remain unfinished.

A.1.1. Sources used that help us for the code

Sources used to help us in the development of the various scripts:

- <https://www.youtube.com/watch?v=rDBho83SUrw> consulted: 09/09/22 at 20h48
- <https://pypi.org/project/webdriver-manager/> consulted: 10/02/22 at 11h12
- <https://selenium-python.readthedocs.io/locating-elements.html#locating-elements> consulted: 10/09/22 at 10h24
- <https://selenium-python.readthedocs.io/waits.html> consulted: 10/09/22 at 10h28
- https://selenium-python.readthedocs.io/api.html#module-selenium.webdriver.support.expected_conditions consulted: 10/09/22 at 11h30
- <https://stackoverflow.com/questions/64425419/how-to-make-sure-selenium-does-not-cl> consulted: 10/09/22 at 11h47
- <https://medium.com/analytics-vidhya/the-art-of-not-getting-blocked-how-i-used-sele> consulted: 19/09/2022 at 20h21
- <https://stackoverflow.com/questions/20986631/how-can-i-scroll-a-web-page-using-sel> consulted: 20/09/2020 at 14h27

A.1.2. Module.py

```
1 from lib2to3.pgen2 import driver
2 from selenium import webdriver
3 from selenium.webdriver.support import expected_conditions as EC
4 from selenium.webdriver.common.by import By
5 from selenium.webdriver.common.keys import Keys
6 from selenium.webdriver.chrome.options import Options
7 from selenium.webdriver.support.wait import WebDriverWait
8 from selenium.webdriver.support.ui import Select
9 from selenium.webdriver.chrome.service import Service as ChromeService
10 from webdriver_manager.chrome import ChromeDriverManager
11 from bs4 import BeautifulSoup
12 import sys
13 from datetime import datetime
14 from time import sleep
```

```
15 import pandas as pd
```

A.1.3. Bot_test.py

```

1 from Modules import *
2
3 def creat_browser():
4
5     print("-----")
6     print("Wait, we're creating the browser ...")
7     print("-----")
8
9     # use the option functionality to disable notifications
10    chrome_options = Options()
11    chrome_options.add_argument("--disable-notifications")
12
13    # Update the driver automatically --> we don't need to always download
14    # the driver for the right chrome version
15    driver =
    ↳ webdriver.Chrome(service=ChromeService(ChromeDriverManager().install()),
    ↳ options= chrome_options)
16    driver.get("https://www.facebook.com/") # open facebook
17    driver.maximize_window()
18
19    print("-----")
20    print("Browser created successfully")
21    print("-----")
22    return driver
23
24 def connect(driver, email, password):
25
26    #accept cookies
27    cookies_accept_button = WebDriverWait(driver,
    ↳ 30).until(EC.element_to_be_clickable((By.XPATH,
    ↳ "//button[@class='_42ft _4jy0 _9xo7 _4jy3 _4jy1 selected _5lsy']")))
28    sleep(2) # Humans behavior (to not go to fast as a bot could be)
29    cookies_accept_button.click()
30    print("-----")

```



```

31     print("Cookies accepted")
32     print("-----")
33
34     print("-----")
35     print("Try to connect with : ")
36     print("  Email : " + email)
37     print("  Password : " + password)
38     print("-----")
39     email_box = driver.find_element(By.ID, "email")
40     password_box = driver.find_element(By.ID, "pass")
41     email_box.send_keys(email)
42     password_box.send_keys(password)
43     sleep(2)
44     login_button = driver.find_element(By.XPATH, "//button[@class='_42ft
↪ _4jy0 _6lth _4jy6 _4jy1 selected _5isy']")
45     login_button.click()
46
47     sleep(15)
48     # check if we logged successfully
49     if driver.title != "Facebook - Connexion ou inscription" :
50         print("-----")
51         print("LOGGED SUCCESSFULLY")
52         print("-----")
53         driver.get("https://www.facebook.com/")
54
55     # A problem occure with the connection
56     else :
57         print("-----")
58         print("/!\ ERROR : We were unable to login... ")
59         print("-----")
60
61 def scroll_header_page(driver, h):
62     try :
63         Cross = WebDriverWait(driver,
↪ 30).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]/div
↪ /div/div/div[3]/div/div[1]/div/div/div/div/div[1]/span[2]")))
64         Cross.click()
65     except :
66         pass

```

```

67 height = WebDriverWait(driver,
↪ 30).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]/div
↪ /div/div/div[2]/div[1]")))).size["height"]
68 h += height + 16 * 3
69 driver.execute_script("window.scrollTo(0, " + str(h) + ");")
70 print(h)
71 sleep(2)
72 height = WebDriverWait(driver,
↪ 30).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div[1]/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[2]")))).size["height"]
73 h += height + 16
74 driver.execute_script("window.scrollTo(0, " + str(h) + ");")
75 print(h)
76
77 sleep(2)
78
79 def scroll(driver, h):
80     nbr_post = 1
81     while nbr_post < 1 :
82         height = WebDriverWait(driver,
↪ 30).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div["+str(nbr_post)+"]"))))
↪ .size["height"]
83         h += height + 16
84         print("post n " + str(nbr_post) + " height " + str(h))
85         nbr_post += 1
86         driver.execute_script("window.scrollTo(0, " + str(h) + ");")
87         # Wait to load the page.
88         sleep(2)
89
90
91 def get_content_post():
92     height = WebDriverWait(driver,
↪ 30).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div["+str(nbr_post)+"]"))))
↪ .size["height"]

```

```

93
94 def get_type_and_number_reactions(self):
95     # Will contain the number of
96     # [like, love, whoaw, hug, lauch, cry, angry]
97     lst_of_reaction = [0, 0, 0, 0, 0, 0, 0]
98     try :
99         reactions_button = WebDriverWait(self.driver,
100     ↪ 5).until(EC.presence_of_element_located((By.XPATH,
101     ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
102     ↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]/div
103     ↪ /div/div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]/div/div
104     ↪ /div/div/div[1]/div/div[1]/div/span/div/span[2]/span/span"))))
105     reactions_button.click()
106     for i in range(3, 7):
107         logo = str(WebDriverWait(self.driver,
108     ↪ 5).until(EC.presence_of_element_located((By.XPATH,
109     ↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]/div/div/div
110     ↪ /div/div/div/div[1]/div/div[1]/div/div/div/div[2]
111     ↪ /div["+str(i)+"]/div/div[1]/img")))).getAttribute("src"))
112
113     # This part is not finish
114
115     except Exception as e:
116         self.selenium_error += 1
117         print("--- ERROR to get the type and number of reactions on the
118     ↪ post ---")
119         print(e)
120
121 if __name__ == '__main__':
122     email = sys.argv[1]
123     password = sys.argv[2]
124     h = 0
125     driver = creat_browser()
126
127     connect(driver, email, password)
128     sleep(2)
129     scroll_header_page(driver, h)
130     h = 424
131     scroll(driver, h)

```

```
122 sleep(500)
```

A.1.4. NewBot.py

```
1
2 from select import select
3 from traceback import print_last
4 from Modules import *
5
6 class NewBot():
7
8     def __init__(self, email, password) -> None:
9         self.height_already_scroll = 0 # The height we have already
10                                         # scrolled
11         self.email = email # The email address of the account we will
12                             # connect with
13         self.password = password # The password of the account with which
14                                   # we will connect.
15         self.driver = None # Driver we'll create to connect to FB
16         self.url = "https://www.facebook.com/" # URL of FB
17         self.nbr_post = 1 # Post on which we are
18         self.fichier = open("data.txt", "w", encoding="utf-8")
19
20     def close(self):
21         """
22         Close the driver.
23         return :
24             - 1 if successfull
25             - 0 if an error occurred
26         """
27         self.fichier.close()
28         try :
29             self.driver.close()
30             return 1
31         except Exception as e:
32             print("--- ERROR to close de driver ---")
33             print(e)
34             return 0
```

```

35
36 def creat_driver(self):
37     """
38     Creat a driver. The driver will use a Google Chrome browser with
39     the options :
40     - --disable-notifications : use to disable the notification pop-up
41       of google chrome
42     - --windows-size=1920,1080 : use to set the size of the windows of
43       google chrome
44     - --disable-gpu : to use less ressources
45
46     This function set the self.driver to the driver created if it has
47     been successfully created
48
49     return :
50         - 1 if the driver is successfully created
51         - 0 if an error occurred
52     """
53
54     try :
55         print("--- Wait, we're creating the browser ... ---")
56         chrome_options = Options()
57         chrome_options.add_argument('--disable-notifications')
58         chrome_options.add_argument('--windows-size=1920,1080')
59         chrome_options.add_argument('--disable-gpu')
60
61         # Creat the driver with the options set previously
62         # Update the driver automatically --> we don't need to always
63         # download the right driver for the right chrome version
64         driver =
↪ webdriver.Chrome(service=ChromeService(ChromeDriverManager().install()),
↪ options= chrome_options, service_log_path='/dev/null')
65
66         driver.get(self.url)      # Got to Facebook
67         driver.maximize_window()  # Maximize the windows on the screen
68         self.driver = driver
69         print("--- Browser created successfully ---")
70         return 1
71
72     except Exception as e:

```

131

```

111
112         sleep(2)
113
114         # Click on the button login
115         login_button = self.driver.find_element(By.XPATH,
116 ↪ "//button[@class='_42ft _4jy0 _6lth _4jy6 _4jy1 selected _51sy']")
117         login_button.click()
118
119         sleep(5) # Wait the page that the page is load
120
121         # Check if we logged successfully according to the title
122         # of the page
123         if self.driver.title != "Facebook - Connexion ou inscription" :
124             print("--- LOGGED SUCCESSFULLY ---")
125             return 1
126
127         # A problem occure with the account
128         else :
129             print("--- ERROR : We were unable to login... ---")
130             print("--- Check that the email and password you give were
131 ↪ correct. ---")
132             return 0
133
134         except Exception as e:
135             print("--- ERROR in the connection ---")
136             print(e)
137             return 0
138
139         # TODO : change the way to refresh the page
140         def refresh_page(self):
141             """
142             Refresh the pages
143
144             Return :
145                 - 1 if it refresh successfully
146                 - 0 if an error occure
147             """
148             try :
149                 self.driver.get("https://www.facebook.com/")
150                 print("--- refresh the page ---")

```

```

    return 1
except Exception as e:
    print("--- ERROR to refresh the page ---")
    print(e)
    return 0

def scroll_header_page(self):
    """
    When we arrive on Facebook, we gave first a field with the stories
    and a field where we can write something to post. So this function
    will scroll those two fields.

    Moreover, sometime there is another field where Facebook ask us
    if we want to stay connected.

    This function press the crosse button of this field to make
    it disappear.

    Return :
        - 1 if it scroll the header successfully
        - 0 if an error occurred
    """

    # First, check if there is the field to stay connected
    try :
        Cross_button = self.driver.find_element(By.XPATH,
→ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
→ /div/div/div/div[3]/div/div[1]/div/div/div/div /div[1]/span[2]")
        Cross_button.click() # close this field
    except :
        pass

    # Second, scroll the two fields (stories and to write something)
    try :

        # 1) Scroll the story field
        # Get the heigh of the boxe
        height = self.driver.find_element(By.XPATH,
→ "//div/div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
→ /div/div/div/div[2]/div[1]").size["height"]

```



```

185         # Calculate the height. 16 = the space between two
186         # posts, Here there is 3 spaces
187         self.height_already_scroll += height + 16*3
188         self.driver.execute_script("window.scrollTo(0, " +
    ↪ str(self.height_already_scroll) + ");") # Scroll
189
190
191         # 2) Scroll the "to write something" field
192         height = self.driver.find_element(By.XPATH,
    ↪ "//div/div[1]/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
    ↪ /div/div/div/div[3]/div/div[2]").size["height"]
193         self.height_already_scroll += height +16
194         self.driver.execute_script("window.scrollTo(0, " +
    ↪ str(self.height_already_scroll) + ");")
195
196     except Exception as e:
197         print("--- ERROR to scroll the header of Facebook ---")
198         print(e)
199         return 0
200
201     print("--- Scroll the header successfully ---")
202     return 1
203
204     def scroll(self):
205         """
206         Scroll the post on which we are. The attribut self.nbr_post say us
207         on which post we are.
208
209         return :
210             - 1 if it scroll the post successfully
211             - 0 if an error occurred
212         """
213
214         try :
215
216             # Get the heigh of the post
217             height = self.driver.find_element(By.XPATH,
    ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
    ↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]")
    ↪ .size["height"]

```

```

218
219         if height == 0: # if the heigh is 0, it means there is no post
220
221             driver.execute_script("window.scrollTo(0,
↪ document.body.scrollHeight);")
222             # Get the heigh of the post
223             height = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]")
↪ .size["height"]
224
225         if height == 0 :
226             print("--- ERROR : the height to scroll is 0 ---")
227             return 0
228
229         self.height_already_scroll += height + 16 # Add the height to
230                                                    # the previous
231                                                    # height scroll
232         self.driver.execute_script("window.scrollTo(0, " +
↪ str(self.height_already_scroll) + ");") # Scroll
233         print("--- Scroll post nbr " + str(self.nbr_post))
234         self.nbr_post += 1 # We scroll another post
235
236     except Exception as e:
237         print("--- ERROR to scroll a post ---")
238         print(e)
239         return 0
240
241     return 1
242
243     def get_number_of_reactions_of_the_post(self):
244         """
245         Get the number of reactions of a post. All type
246         combined (like, love, hate, ...).
247
248         So if there is 10 likes, 5 loves and 6 hates. We get the number 21.
249
250         Return : the number of reactions
251         """
252         try :

```

```

253         # Get the reaction field of the current post on which we are
254         reactions = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]/div/
↪ div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]/div/div/
↪ div/div/div[1]/div/div[1]/div/span/div/span[2]/span/span")
↪ .get_attribute("innerHTML")
255         return reactions
256     except :
257         # If we enter here, it's means we didn't find the reactions
258         # field and therefore there is no reaction.
259         return 0
260
261     def get_all_reactions(self):
262         """
263         Get the different types of reactions there is on a post and the
264         number of each ones.
265
266         For example, if there is 10 likes, 5 loves and 6 hates, we'll
267         consider each number individually and no combine them as in
268         the function get_number_of_reactions_of_the_post(self).
269
270         The function use a list of size 7 which will contain the number
271         of reactions in this order :
272         [like, love, whoaw, hug, lauch, cry, angry]
273
274         Here are the step the function will do to get those number :
275         1) Click on the reaction button. A pop-up will appears.
276         2) Get the number of reaction of the different types.
277         3) If there is more than 4 differents type of reaction, we
278         need to click on the "plus" button to get the other reactions.
279         4) Get the other reactions if there are.
280         5) Close the pop-up by clicking on the cross button
281
282         Return :
283         - A list of size 7 which will contain the number of reaction in
284         this order : [like, love, whoaw, hug, lauch, cry, angry]
285         - 0 if an error occurred
286         """
287

```

```
# The pictures of the reactions use by Facebook.
# I have formatted them because sometime the beginning
# and the end of the URL change from a connection to another.
# So I removed the beginning (http://...) and the end of the
# URL (after the "?" symbole in the original URL).
like = "An-tsvy1nZCAfjJDq_e9hwhgJ_ouDg6GOHdVQtc31Lh3B1
→ 3GEFJON3wRI6j2_Lz8icCyU4RkVsKbJckG5NMDv5TxxWie80qB_kcv
→ CNizVjn7sw.png"
waouh = "An_f5KryEO3JdbkuRbEs1ixj8HC8itKTXvZ3Hl1czaREaiMDPCRTN
→ w6CSwRUjKkq_YXEuxmsqBu06WIeteZ7MBZ2WkuJXvOK6Wd0QfGi2Ixcg9Sd.png"
love = "An8KxJw0TdKA0hIHqkw35xWvBGYLLbtgD5y14_K8iN_zaDhCWgixkt
→ WzvqA45BTxHACGktnPMx_lkq1uE66153QNE58NZp59iYz6MDdtqgcTZw.png"
cry = "An9Yzzh8CoEGqeWIfY5w6zR3VdPbG5X1fHXZdMfftnoomx30bysBj14
→ 5G99ZhM1T6DcU_ZAH2bEdi0j8sUAQvplVo0cYKS_GprBBJlcwiBHomFx7hQ.png"
laugh = "An9yRlv3tqyIsDTiKV0WfMgtabNG9VPyvNiv5USdzPe0Cb2FdnMvbg
→ H1mvTvI8TczUcd9kED-M5Q1z9-fVK3zAMCRSiYtsWTPwSid0DJlPasg.png"
hug = "An-zbifsrGonJXBikRGgj1txFJUkRG3aCN5900mzKo6dVL8tKCuwWf6
→ D90v6XB3JJZ7pT1FSxuFsET0jkjZ08b5AyPU0z_GsxZH2nWiW2ScJT4p3rQ.png"
angry = "An-mjOuPEZ5b6GVy30C-_ZMV1AGoboZI3SG9P2r3WElt0540lpAmUSq
→ 9QPU0i9RdhF07UwCRHIsC06i-w4_VCrnJnBEent1vmcy8MX0Qt0msew.png"

# Wreat the list of size 7 which will contain the number of
# reaction in this order :
# [like, love, whoaw, hug, laugh, cry, angry]
lst_of_reaction = [0, 0, 0, 0, 0, 0, 0]

WebDriverWait(self.driver,
→ 10).until(EC.presence_of_element_located((By.XPATH,
→ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]/div/div/div
→ /div/div/div/div[1]/div/div[1]/div/div/div/div[2]/div["+str(2)+"]/"
→ div/div[1]"))))

is_more_than_one_reaction = False
for i in range(3, 10):
    # Get the differents reactions and numbers
    try :
        # Picture of the reaction
        logo = str(self.driver.find_element(By.XPATH,
→ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]/div/div/"
→ div/div/div/div/div[1]/div/div[1]/div/div/div/div[2]/div["+str(i)+"]/
```

```

    ↪ /div/div[1]/img").get_attribute("src")).split("?")[0]
314         logo = logo.split("/")
315         logo = logo[len(logo)-1]
316         # Number of the reaction
317         number = self.driver.find_element(By.XPATH,
    ↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]/div/div/
    ↪ div/div/div/div/div[1]/div/div[1]/div/div/div/div[2]/div["+str(i)+"
    ↪ /div[1]/span").get_attribute("innerHTML")
318         is_more_than_one_reaction = True
319         # Find the type of reaction and add the number in the
320         # list "lst_of_reaction"
321         if logo == like :
322             lst_of_reaction[0] = number
323         elif logo == love:
324             lst_of_reaction[1] = number
325         elif logo == waouh:
326             lst_of_reaction[2] = number
327         elif logo == hug :
328             lst_of_reaction[3] = number
329         elif logo == laugh :
330             lst_of_reaction[4] = number
331         elif logo == cry :
332             lst_of_reaction[5] = number
333         elif logo == angry :
334             lst_of_reaction[6] = number
335
336     except :
337         # If there is an exeption, either it means there is no more
338         # reaction or there is only one type of reaction
339         # Test if there is only one type of reaction
340         if not is_more_than_one_reaction :
341             # Picture of the reaction
342             logo = str(self.driver.find_element(By.XPATH,
    ↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]
    ↪ /div/div/div/div/div[1]/div/div[1]/div/div/div/div[2]
    ↪ /div["+str(2)+"]/div/div[1]/img").get_attribute("src")).split("?")[0]
343             logo = logo.split("/")
344             logo = logo[len(logo)-1]
345             # Number of the reaction

```

```

346         number = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]
↪ /div/div/div/div/div/div/div[1]/div/div[1]/div/div/div/div[2]
↪ /div["+str(2)+"]/div[1]/span/span").get_attribute("innerHTML")
347         # Find the type of reaction and add the number in the
348         # list "lst_of_reaction"
349         if logo == like :
350             lst_of_reaction[0] = number
351         elif logo == love:
352             lst_of_reaction[1] = number
353         elif logo == waouh:
354             lst_of_reaction[2] = number
355         elif logo == hug :
356             lst_of_reaction[3] = number
357         elif logo == laugh :
358             lst_of_reaction[4] = number
359         elif logo == cry :
360             lst_of_reaction[5] = number
361         elif logo == angry :
362             lst_of_reaction[6] = number
363         break
364     try :
365         # Try to click on the plus button to get the other reactions
366         plus_button = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]
↪ /div/div/div/div/div/div/div[1]/div/div[1]/div/div/div/
↪ div[2]/div[1]/div/div[1]/span")
367         plus_button.click()
368         for i in range(1,5):
369             # Get the differents reactions and numbers
370             try :
371                 # Picture of the reaction
372                 logo = str(self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[3]
↪ /div/div/div[1]/div[1]/div/div/div/div/div/div/div[1]/div/div[" +
↪ str(i) + "]" /div[1]/img").get_attribute("src")).split("?")[0]
373                 logo = logo.split("/")
374                 logo = logo[len(logo)-1]
375                 # Number of the reaction

```

```

376         number = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[3]
↪ /div/div/div[1]/div[1]/div/div/div/div/div/div/div[1]/div/div[" +
↪ str(i) + "]" /div[2]/div/div/span").get_attribute("innerHTML")
377
378         # Find the type of reaction and add the number in the
379         # list "lst_of_reaction"
380         if logo == like :
381             lst_of_reaction[0] = number
382         elif logo == love :
383             lst_of_reaction[1] = number
384         elif logo == waouh :
385             lst_of_reaction[2] = number
386         elif logo == hug :
387             lst_of_reaction[3] = number
388         elif logo == laugh :
389             lst_of_reaction[4] = number
390         elif logo == cry :
391             lst_of_reaction[5] = number
392         elif logo == angry :
393             lst_of_reaction[6] = number
394     except :
395         # If there and exeption, it means there is no
396         # more reaction
397         break
398 except :
399     # If there is an exception, it means there is no "plus" button.
400     # Therefore, we've already got all the reaction
401     pass
402
403 try :
404     # close the pop-up
405     close_button = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[4]/div/div/div[1]/div/div[2]
↪ /div/div/div/div/div/div/div[1]/div/div[2]/div")
406     close_button.click()
407
408 except Exception as e:
409     print("--- ERROR to get the cross button ---")
410     print(e)

```

```

411         return 0
412
413         return lst_of_reaction
414     def get_types_and_number_of_reactions_of_the_post(self):
415
416         lst_of_reaction = [0,0,0,0,0,0,0]
417         try :
418             reactions_button = self.driver.find_element(By.XPATH,
419 ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]
420 ↪ /div/div[2]/div/div/div/div[3]/div/div[3]/div
421 ↪ /div["+str(self.nbr_post)+"]/div/div/div/div/div/div/div/div
422 ↪ /div/div/div/div[2]/div/div[4]/div/div/div/div
423 ↪ /div[1]/div/div[1]/div/span/div/ span[2]/span/span")
424             reactions_button.click()
425         except :
426             try :
427                 # Sometime the XPATH is different, so we try with
428                 # another XPATH
429                 reactions_button = self.driver.find_element(By.XPATH,
430 ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]
431 ↪ /div[1]/div/div[2]/div/div/div/div[3]/div/div[3]/div
432 ↪ /div["+str(self.nbr_post)+"]/div/div/div/div/div/div/div/div
433 ↪ /div/div/div/div[2]/div/div[4]/div/div/div/div/div[1]/div/div[1]/div
434 ↪ /span/div/span[1]/span/span")
435                 reactions_button.click()
436             except :
437                 # If we didn't find the reaction button, it's mean there is
438                 # no reactions button
439                 return lst_of_reaction
440         return self.get_all_reactions()
441
442     def get_if_it_is_a_recommanded_post(self):
443         """
444         Get if this is a recommended post by Facebook
445
446         return :
447             - "It's a recommanded post" if it is
448             - "It's NOT a recommanded post" if not
449         """
450         try :

```



```

441         # Try to get the field
442         recommended = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]/div
↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[1]/div/div
↪ /div/div/span").get_attribute("innerHTML")
443         return "It's a recommended post"
444     except :
445         # If we don't get it, it means it's not a recommended post
446         return "It's NOT a recommended post"
447
448     def get_author_of_the_post(self):
449         """
450         Get the author of a post.
451
452         return :
453             - the author if no errors occurred
454             - 0 otherwise
455         """
456
457         try :
458             # get the author name
459             author = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[2]
↪ /div/div[2]/div/div[1]/span/h4/span/a/strong/span")
↪ .get_attribute("innerHTML")
460             return author
461
462         except Exception as e1:
463             # ERROR first try to get the author of a post
464             # Sometime, the XPATH is different
465             try :
466                 # try with another Xpath for the author
467                 author = self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[2]/div
↪ /div[2]/div/div[1]/span/h4/span/strong[1]/span/a/span/span")

```

```

468     ↪ .get_attribute("innerHTML")
469         return author
470
471     except Exception as e2 :
472         print("--- ERROR to get the author of the post ---")
473         print("--- Error first try : ---")
474         print(e1)
475         print("--- Error second try : ---")
476         print(e2)
477         return 0
478
479     def get_date_of_the_post(self):
480         """
481         Get the date at which the post has been made
482
483         return :
484             - The date if no error occurred
485             - 0 otherwise
486         """
487         try :
488             # Get the date
489             date = self.driver.find_element(By.XPATH,
490 ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
491 ↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
492 ↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[2]/div
493 ↪ /div[2]/div/div[2]/span/span/span[2]/span/a/span")
494             ↪ .get_attribute("innerHTML")
495             return date
496
497         except Exception as e1:
498             # ERROR first try to get the date of the post
499             # Sometime, the XPATH is different
500             try :
501                 # try with another Xpath for the date
502                 date = self.driver.find_element(By.XPATH,
503 ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
504 ↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
505 ↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[2]
506 ↪ /div/div[2]/div/div[2]/span/span/span[2]/span")

```

```

498         .get_attribute("innerHTML")
499         return date
500     except Exception as e2 :
501         print("--- ERROR second try to get the date of the post
502         ↪ ---")
503         print("--- Error first try : ---")
504         print(e1)
505         print("--- Error second try : ---")
506         print(e2)
507         return 0
508
509 def get_number_of_comments_of_the_post(self):
510     """
511     Get the number of a comments of a post
512
513     return :
514         - the number of comments
515     """
516     try :
517         comments = self.driver.find_element(By.XPATH,
518         ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
519         ↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
520         ↪ /div/div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]
521         ↪ /div/div/div/div/div/div[1]/div/div[2]/div[2]/span/div/span")
522         .get_attribute("innerHTML")
523         return comments
524     except :
525         # If there is an exception, it means there is not the comments
526         # field and therefor there is no comment.
527         return 0
528
529 def get_number_of_shares_of_the_post(self):
530     """
531     Get the number of shares of a post
532
533     return :
534         - The number of shares
535     """
536     try :

```

```

530         shares = self.driver.find_element(By.XPATH,
↪     "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪     /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
↪     /div/div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]/div
↪     /div/div/div/div[1]/div/div[2]/div[3]/span/div/span")
↪     .get_attribute("innerHTML")
531         return shares
532
533     except:
534         # ERROR first try to get the number of shares on the post
535         try :
536             # Try with another Xpath for the date
537             # Sometime, the XPATH is different
538             shares = self.driver.find_element(By.XPATH,
↪     "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪     /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
↪     /div/div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]
↪     /div/div/div/div/div[1]/div/div[3]/div[2]/span/div/span")
↪     .get_attribute("innerHTML")
539             return shares
540         except :
541             # If there is an exception, it means there is not the field
542             # of share and therefore there is no share
543             return 0
544
545     def get_description_of_the_post(self):
546         """
547         Get the description of a post
548
549         return :
550             - The description if there is one.
551             - 0 otherwise
552         """
553         try :
554             # get the description
555             description = self.driver.find_element(By.XPATH,
↪     "/html/body/div/div/div[1]/div/div[3]/div/div/div/div[1]
↪     /div[1]/div/div[2]/div/div/div/div[3]/div/div[3]/div/div["+
↪     str(self.nbr_post)+"]/div/div/div/div/div/div/div/div
↪     /div/div/div/div[2]/div/div[3]/div[1]/div/div/div

```

```

556     ↪ /span/div/div").get_attribute("innerHTML")
557         return description
558
559     except :
560         # if there is an exception, it means there is no
561         # description field.
562         return 0
563
564     def get_information_of_a_post(self):
565         """
566         Get all the information of a post excpet the comments.
567         So it get :
568         - the author
569         - the date at which the post has been posted
570         - the number of reactions
571         - the number of the differents type of reactions
572         - the number of comments
573         - the number of shares
574         - the descriptions of the post
575         - the content of the post (video, picture, text, ...)
576
577         return :
578         - a JSON object with all those informations.
579         - 1 if it is a reels fields
580         - 0 if an error occurred
581         TODO : create the JSON object
582
583         """
584
585         # Test if it's a field with reels and short videos. We don't treat
586         ↪ reels posts
587         try :
588             reel = self.driver.find_element(By.XPATH,
589             ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
590             ↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
591             ↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div[1]/div
592             ↪ /div[1]/div/div/div[1]/div/span").get_attribute("innerHTML")
593
594             if reel == "Reels et vid os courtes":
595                 # If we reach here, then the post is reels and short
596                 # videos field.

```

```

590         print("--- Skip Post : it's a reels and short videod ---")
591         return 1
592     except :
593         pass
594
595     # GET IF IT'S A RECOMMENDED POST
596     recommended = self.get_if_it_is_a_recommended_post()
597
598     # GET THE AUTHOR OF THE POST
599     author = self.get_author_of_the_post()
600     if author == 0:
601         return 0
602
603     # GET THE DATE OF THE POST
604     date = self.get_date_of_the_post()
605     if date == 0 :
606         return 0
607
608     # GET THE NUMBER OF REACTIONS
609     reactions = self.get_number_of_reactions_of_the_post()
610
611     # GET THE TYPE AND NUMBER OF REACTIONS
612     lst_of_reaction =
    ↪ self.get_types_and_number_of_reactions_of_the_post()
613
614     # GET THE NUMBER OF COMMENTS
615     comments = self.get_number_of_comments_of_the_post()
616
617     # GET THE NUMBER OF share
618     shares = self.get_number_of_shares_of_the_post()
619
620     # GET THE DESCRIPTION
621     description = self.get_description_of_the_post()
622
623     self.fichier.write("\n" + recommended)
624     self.fichier.write("\n" + str(author))
625     self.fichier.write("\n" + str(date))
626     self.fichier.write("\n" + str(reactions))
627     self.fichier.write("\n" + str(lst_of_reaction))
628     self.fichier.write("\n" + str(comments))

```

```

629         self.fichier.write("\n" + str(shares))
630         self.fichier.write("\n" + str(description))
631         print(recommended)
632         print("Author : " + str(author))
633         print("Date : " + str(date))
634         print("Nbr of reactions : " + str(reactions))
635         print("list of reactions : "+ str(lst_of_reaction))
636         print("Nbr of comments : " + str(comments))
637         print("Nbr of shares : " + str(shares))
638         print("Description : " + str(description))
639
640         return 1 # TODO : change to a JSON object
641
642     def get_the_comments(self):
643         """
644         Get the comments of a post
645
646         return :
647             - the number of comments
648         """
649
650         #TODO :
651         # - get autor, date, comment, number of reply, reactions
652         try :
653             # Open the comments pop-up
654             comments_button = WebDriverWait(self.driver,
655             ↪ 2).until(EC.presence_of_element_located((By.XPATH,
656             ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
657             ↪ /div/div/div/div[3]/div/div[3]/div/div["+str(self.nbr_post)+"]
658             ↪ /div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]/div
659             ↪ /div/div/div/div[1]/div/div[2]/div[2]/span/div/span")))
660             comments_button.click()
661
662             # This command is use to do a "sleep": when it find the
663             ↪ element,
664             # it means the comments are loaded
665             WebDriverWait(self.driver,
666             ↪ 30).until(EC.presence_of_element_located((By.XPATH,
667             ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]
668             ↪ /div[1]/div/div[2]/div/div/div/div[3]/div/div[3]/div

```

```

↪ /div["+str(self.nbr_post)+"]/div/div/div/div/div/div/div/div
↪ /div/div/div/div[2]/div/div[4]/div/div/div[2]
↪ /div[2]/div/div/div/span"))

660
661     for i in range(2) :
662         number_comment = 1 # On which comment we are
663
664         # First iteration : get all "relevant" comments according
665         # to Facebook
666         if i == 0:
667             type_comment_button =
↪ self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]
↪ ]/div[1]/div/div[2]/div/div/div/div[3]/div/div[3]
↪ /div/div["+str(self.nbr_post)+"]/div/div/div/div/div/div/div/div/div
↪ /div/div /div[2]/div/div[4]/div/div/div[2]/div[2]/div/div /div/span")
668             type_comment_button.click()
669
670             try :
671                 most_relevant_comments_button =
↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
↪ /div/div/div/div[2]/div/div/div[1]/div[1]/div/div/div/div
↪ /div/div/div[1]/div/div[1]/div[1]/div/div[1]/span")
672             except :
673                 try:
674                     most_relevant_comments_button =
↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
↪ /div/div/div/div[2]/div/div/div[1]/div[1]/div/div/div/div/div/div
↪ /div[1]/div/div[1]/div[1]/div/div[2]/span")
675                 except :
676                     print("--- ERROR to get the most relevant
↪ comments button ---")
677
678             most_relevant_comments_button.click()
679
680             # This command is use to do a "sleep": when it find the
681             # element, it means the comments are loaded
682             WebDriverWait(self.driver,
↪ 30).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div

```



```

        ↪ /div[2]/div/div/div/div[3]/div/div[3]/div/div[" + str(self.nbr_post)
        ↪ + "] /div/div/div/div/div/div/div/div/div/div/div/div[2]
        ↪ /div/div[4]/div/div/div[2]/ul/li[" + str(number_comment) + "]]))
683         print()
684         print("--- Let's get the most relevant comments ---")
685
686         # Second iteration : get all the comments
687         else :
688             # return to the top of the post
689             self.driver.execute_script("window.scrollTo(0, " +
        ↪ str(self.height_already_scroll) + ");") # Scroll
690
691             type_comment_button =
        ↪ self.driver.find_element(By.XPATH,
        ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]
        ↪ /div[1]/div/div[2]/div/div/div/div[3]/div/div[3]/div
        ↪ /div["+str(self.nbr_post)+"]/div/div/div/div/div/div/div
        ↪ /div/div/div/div/div[2]/div/div[4]
        ↪ /div/div/div[2]/div[2]/div/div/div/span")
692             type_comment_button.click()
693
694             try :
695                 all_comments_button =
        ↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
        ↪ /div/div/div/div[2]/div/div/div[1]/div[1]/div/div/div/div
        ↪ /div/div[1]/div/div[3]/div[1]/div/div[1]/span")
696             except :
697                 try:
698                     all_comments_button =
        ↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
        ↪ /div/div/div/div[2]/div/div/div[1]/div[1]/div/div/div/div
        ↪ /div/div[1]/div/div[3]/div[1]/div/div[2]/span")
699             except :
700                 print("--- ERROR to get the all comments
        ↪ button ---")
701
702             all_comments_button.click()
703
704             # This command is use to do a "sleep": when it find the
705             # element, it means the comments are loaded

```

```

706         WebDriverWait(self.driver,
    ↪ 30).until(EC.presence_of_element_located((By.XPATH,
    ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]
    ↪ /div[1]/div/div[2]/div/div/div/div[3]/div/div[3]/div /div[" +
    ↪ str(self.nbr_post) + "]/div/div/div/div/div/div
    ↪ /div/div/div/div/div/div[2]/div/div[4]/div/div/div[2] /ul/li[" +
    ↪ str(number_comment) + "]))))

707
708         print()
709         print("--- Let's get all comments ---")
710
711         still_comment = True
712         while still_comment :
713             # Try 4 differents XPATH. If the 4 fails, it means we
714             # miss a possible Xpath
715             try :
716                 author = self.driver.find_element(By.XPATH,
    ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]
    ↪ /div/div[2]/div/div/div/div[3]/div/div[3]/div/div[" +
    ↪ str(self.nbr_post) + "]/div/div/div/div/div/div
    ↪ v/div/div/div/div/div/div/div/div[2]/div/div[4]/div/div/div[2]/ul /li["
    ↪ + str(number_comment) + "]/div[1]/div[2]/div[2]/div
    ↪ /div[1]/div/div[1]/div/div/span/a/span/span")
    ↪ .get_attribute("innerHTML")
717             except :
718                 try :
719                     author = self.driver.find_element(By.XPATH,
    ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]
    ↪ /div/div[2]/div/div/div/div[3]/div/div[3]/div/div[" +
    ↪ str(self.nbr_post) + "]/div/div/div/div/div/div/div
    ↪ /div/div/div/div/div/div[2] /div/div[4]/div/div /div[2]/ul/li[" +
    ↪ str(number_comment) + "]/div[1]/div/div[2]/div
    ↪ /div[1]/div/div[1]/div/div/span/a/span/span")
    ↪ .get_attribute("innerHTML")
720             except :
721                 try :
722                     author =
    ↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
    ↪ /div/div/div/div[1]/div[1]/div/div[2]/div/div/div/div[3]/div
    ↪ /div[3]/div /div[" + str(self.nbr_post) + "]/div/div/div/div/div/div

```

```

    ↪ /div/div/div/div/div/div[2]/div/div[4]/div /div/div[2]/ul/li[" +
    ↪ str(number_comment) + "]/div[1]/div[2]/div[2]/div[1]
    ↪ /div[1]/div/div/span/a/span/span") .get_attribute("innerHTML")
723         except:
724             try :
725                 author =
    ↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
    ↪ /div/div/div/div[1]/div[1]/div/div[2]/div/div/div/div[3]/div/div[3]
    ↪ /div/div[" + str(self.nbr_post) + "]/div/div/div/div/div/div/div
    ↪ /div/div/div/div[2]/div/div[4]/div/div/div[2]/ul/li[" +
    ↪ str(number_comment) + "]"
    ↪ /div[1]/div/div[2]/div/div[1]/div/div[1]/span/a/span/span")
    ↪ .get_attribute("innerHTML")
726         except :
727             try :
728                 author =
    ↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[1]
    ↪ /div/div/div[2]/div/div/div/div[3]/div/div[3]/div/div[" +
    ↪ str(self.nbr_post) + "]/div/div/div/div/div/div/div
    ↪ /div/div/div/div[2]/div/div[4]/div /div/div[2]/ul/li[" +
    ↪ str(number_comment) + "]/div[1]/div[2]/div[2]/div/div[1]/div/div[1]
    ↪ /div/div/span/a/span/span").get_attribute("innerHTML")
729         except:
730             print("--- Error to get the author
    ↪ of the comment ---")
731             return 0
732
733         # Try 3 differents XPATH. If the 3 fails, it means we
734         # miss a possible Xpath
735         try :
736             date = self.driver.find_element(By.XPATH,
    ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
    ↪ /div/div/div/div[3]/div/div[3]/div/div[" + str(self.nbr_post) + "]"
    ↪ /div/div/div/div/div/div/div/div/div/div/div/div[2]
    ↪ /div/div[4]/div/div/div[2]/ul/li[" + str(number_comment) + "]"
    ↪ /div[1]/div[2]/div[2]/ul/li[3]/a/span").get_attribute("innerHTML")
737         except:
738             try :
739                 date = self.driver.find_element(By.XPATH,
    ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]

```

```

740     ↪ /div/div/div/div[3]/div/div[3]/div/div[" + str(self.nbr_post) + "]
741     ↪ /div/div/div/div/div/div/div/div/div/div/div /div[2]
742     ↪ /div/div[4]/div/div/div[2]/ul /li[" + str(number_comment) +
    except:
        try :
            date = self.driver.find_element(By.XPATH,
743     ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]
744     ↪ /div/div[2]/div/div/div/div[3]/div/div[3]/div /div[" +
    ↪ str(self.nbr_post) + "]/div/div/div/div/div/div/div
    ↪ /div/div/div/div/div[2]/div/div[4]/div/div/div[2]/ul /li[" +
    ↪ str(number_comment) + "]/div[1]/div/div[2]/ul/li[4]/a/span")
    ↪ .get_attribute("innerHTML")
745     except:
746         print("--- Error to get the date of the
    ↪ comment ---")
747         return 0
748
749     # Try 4 differents XPATH. If the 4 fails, it means we
750     # miss a possible Xpath
    try :
        comment_text = self.driver.find_element(By.XPATH,
751     ↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]
752     ↪ /div/div[2]/div/div/div/div[3]/div/div[3]/div /div[" +
    ↪ str(self.nbr_post) + "]/div/div/div/div/div/div/div
    ↪ /div/div/div/div/div[2]/div/div[4]/div/div/div[2] /ul/li[" +
    ↪ str(number_comment) + "]/div[1]/div[2]/div[2]/div
    ↪ /div[1]/div/div[1]/div/div/div/span/div/div")
    ↪ .get_attribute("innerHTML")
753     except:
754         try :
            comment_text =
    ↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
    ↪ /div/div/div/div[1]/div[1]/div/div[2]/div/div/div
    ↪ /div[3]/div/div[3]/div/div[" + str(self.nbr_post) + "]/div
    ↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]
    ↪ /div/div[4]/div/div/div[2]/ul/li[" + str(number_comment) +
    ↪ "]/div[1]/div/div[2]/div/div[1]/div/div[1]
    ↪ /div/div/div/span/div/div").get_attribute("innerHTML")
    except:

```

```

755         try :
756             comment_text =
↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
↪ /div/div/div/div[1]/div[1]/div/div[2]/div/div/div/div[3]/div
↪ /div[3]/div/div[" + str(self.nbr_post) + "]/div/div/div
↪ /div/div/div/div/div/div/div/div/div[2]/div/div[4]/div/div
↪ /div[2]/ul/li[" + str(number_comment) + "]/div[1]/div[2]/div[2]
↪ /div/div[1]/div/div[1]/div/span/div/div").get_attribute("innerHTML")
757         except:
758             try :
759                 comment_text =
↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
↪ /div/div/div/div[1]/div[1]/div/div[2]/div/div/div/div[3]
↪ /div/div[3]/div/div[" + str(self.nbr_post) + "]
↪ /div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]
↪ /div/div/div[2]/ul/li[" + str(number_comment) + "]
↪ /div[1]/div/div[2]/div[1]/div[1]/div/div/div/span/div/div")
↪ .get_attribute("innerHTML")
760             except :
761                 try :
762                     comment_text =
↪ self.driver.find_element(By.XPATH, "//div/div/div[1]/div/div[3]
↪ /div/div/div/div[1]/div[1]/div/div[2]/div/div/div/div[3]/div/div[3]
↪ /div/div[" + str(self.nbr_post) + "]/div/div/div/div/div/div/div
↪ /div/div/div/div/div/div/div[2]/div/div[4]/div/div/div[2] /ul/li[" +
↪ str(number_comment) + "]/div[1]/div
↪ /div[2]/div/div[1]/div/div/div/div/div/div/span/div
↪ /span/div/div/span/img").get_attribute("innerHTML")
763
764             except:
765                 print("--- Error to get the text
↪ of the comment ---")
766
767                 return 0
768
769         self.fichier.write("\n" + "COMMENT n " +
↪ str(number_comment) + " :")
770         self.fichier.write("\n" + str(author))
771         self.fichier.write("\n" + str(date))
772         self.fichier.write("\n" + str(comment_text))

```

```

773         self.fichier.write("\n" + str(reactions))
774         print()
775         print("COMMENT n    " + str(number_comment) + " :")
776         print()
777         print(author)
778         print(date)
779         print(comment_text)
780         print(reactions)
781
782         number_comment += 1
783         try :
784             # Test if there is other comments
785             self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]/div[1]/div/div[2]
↪ /div/div/div/div[3]/div/div[3]/div/div[" + str(self.nbr_post) + "]"
↪ /div/div/div/div/div/div/div/div/div/div/div/div/div/div[2]/div/div[4]
↪ /div/div/div[2]/ul/li[" + str(number_comment) + "]")
786         except :
787             # If we arrive here, it means we didn't find
788             # the next element.
789             # 1) either because he is not already loaded
790             # 2) either because there is no more comment
791
792             # 1) test to load more comments :
793             try :
794
795                 # Get the actual height of the the post (post +
796                 # comment space) to scroll to the next comment
797                 # (if there is) after generated them
798                 try :
799                     scroll = self.height_already_scroll
800                     height =
↪ self.driver.find_element(By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]
↪ /div[1]/div/div[2]/div/div/div/div[3]/div/div[3] /div/div[" +
↪ str(self.nbr_post) + "]").size["height"]
801                     scroll += height
802             except :
803                 pass
804

```

```

805         get_other_comments_button =
↪ WebDriverWait(self.driver,
↪ 3).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div/div[1]
↪ /div[1]/div/div[2]/div/div/div/div[3]/div/div[3]/div/div[" +
↪ str(self.nbr_post) + "]/div/div/div/div/div
↪ /div/div/div/div/div/div/div[2]/div/div[4]/div
↪ /div/div[2]/div[4]/div[1]/div[2]/span/span"))))
806         get_other_comments_button.click()
807
808         WebDriverWait(self.driver,
↪ 30).until(EC.presence_of_element_located((By.XPATH,
↪ "//div/div/div[1]/div/div[3]/div/div/div
↪ /div[1]/div[1]/div/div[2]/div/div/div/div[3]/div/div[3]/div /div[" +
↪ str(self.nbr_post) + "]/div/div/div/div/div/div/div/div/div/div/div
↪ /div[2]/div/div[4]/div/div/div[2]/ul /li[" + str(number_comment) +
↪ "]""))))
809         self.driver.execute_script("window.scrollTo(0,
↪ " + str(scroll) + ");") # Scroll
810
811         except :
812             print()
813             print("--- No more comment ---")
814             still_comment = False
815
816         except :
817             # If there is an exception, it means there is not the comments
818             # field and therefor there is no comment.
819             return 0

```

A.1.5. Script_Navigation_bot.py

```

1 from NewBot import *
2
3 if __name__ == '__main__':
4     email = sys.argv[1]
5     password = sys.argv[2]
6

```

```

7 bot = NewBot(email, password)
8 return_value = bot.creat_driver()
9
10 if return_value == 1:
11     return_value = bot.connect()
12 else :
13     bot.close()
14
15 if return_value == 1 :
16     bot.refresh_page()
17 else :
18     bot.close()
19
20 if return_value == 1 :
21     bot.scroll_header_page()
22 else :
23     bot.close()
24
25 for i in range(5):
26     if return_value == 1 :
27         bot.get_information_of_a_post()
28         bot.get_the_comments()
29         return_value= bot.scroll()
30         sleep(2)
31     else :
32         return_value = bot.close()
33     break

```

A.2. Codes for the second experiment

This section consists of two parts. Firstly, the files for the browser extension, and secondly, the Python script that extracts information from an HTML file. The browser extension is fully functional, which is not the case for the Python script. Indeed, we realized that this technique would not work, so we stopped coding.

A.2.1. browser extension

The browser extension is divided into three files: `manifest.json`, `get_html.js` and `get_html_script.js`.

When you add the extension to your browser, the `manifest.json` file triggers the execution of the script `get_html.js` in the background. When you click on the extension logo, `get_html.js` calls `get_html_script.js`, which prompts you to enter a name for the HTML file of the current page. The file is then downloaded with a name consisting of the timestamp at the moment you clicked the extension, followed by a hyphen (-), and the name you entered.

The reason why the `get_html.js` script calls the `get_html_script.js` script instead of including everything in `get_html.js` is because of the execution environment. Initially, we had all the code in `get_html.js`, but we encountered an issue where the HTML file obtained did not contain the page's HTML. After investigation, we discovered that `get_html.js` was running in the background and not in the same environment as the current page. To overcome this, we created the `get_html_script.js` script, which runs in the same environment as the page we are on, allowing us to retrieve the desired HTML content successfully.

To add the extension to your browser (specifically for Firefox), follow these steps:

1. Open your browser and go to the extensions section.
2. In the *Module Debugger* (or *Debugging* or similar) tab, click on *Load Temporary Add-on...* (or similar option).
3. Select the `manifest.json` file from the directory where you have all the extension files located. (Note: Make sure all the required files for the extension are present in the same directory.)
4. The extension will be loaded temporarily in your browser, and you can now use its functionality.

Please note that the exact steps may vary slightly depending on the browser you are using. In the same directory, you should also include an image named `icon.png` which will serve as the logo for the extension once it is added.

A.2.1.1. manifest.json

The `default_icon` attribute is used to put an image to your extension when it is displayed in your browser.

```
1 {
2   "manifest_version": 2,
3   "name": "Get the html",
4   "description": "Description",
5   "version": "1.0",
6   "permissions": ["downloads", "activeTab"],
7   "browser_action": {
8     "default_icon": "icon.png"
9   },
10  "background": {
11    "scripts": ["get_html.js"]
12  }
13 }
```

A.2.1.2. get_html.js

```
1 chrome.browserAction.onClicked.addListener(function(tab) {
2     chrome.tabs.executeScript(null, {file: "get_html_script.js"});
3 });
```

A.2.1.3. get_html_script.js

```
1 var html = document.body.outerHTML;
2
3
4 function download(text){
5     var blob = new Blob([text], {type: "text/plain"});
6     var url = window.URL.createObjectURL(blob);
7     var a = document.createElement("a");
8     a.href = url;
```

```

9      const currentDate = new Date();
10     const timestamp = currentDate.getTime();
11     var filename = prompt('Entrez le nom du fichier :');
12     a.download = timestamp.toString().concat("-", filename, ".html");
13     a.click();
14 }
15
16 download(html);

```

A.2.2. Python script

This file employs the same technique as the Python script from the first experiment. It utilizes the Selenium module but instead of opening a browser, it directly uses an HTML file (downloaded with the extension) to extract various information from each Facebook post.

However, as mentioned before, we realized that this approach would not be successful, so we halted the development of the script. Therefore, the code remains incomplete.

```

1
2 from selenium import webdriver
3 import csv
4 import os
5 from os import listdir
6 from selenium.webdriver.firefox.options import Options as FirefoxOptions
7 from selenium.webdriver.common.by import By
8
9 from bs4 import BeautifulSoup
10
11 from datetime import date
12 import datetime
13 import time
14
15 # First part of the Xpath to reach the different posts. This part varies
16     ↳ depending on the computer. So everyone who runs this script must
17     ↳ configure it with their Xpath.
18 xpath = "div[1]/div/div[1]/div/div[3]/div/div/div/div/div/div
19     ↳ /div[2]/div/div/div/div[3]/div/div[3]/div/div"
20
21 def creatFinalCsv(name_file):

```

```

19     rows = ["Sponsored", "Author", "Time", "Suggestion", "Description",
↪ "Like", "Comment", "Share", "Href"]
20     dir_data = "./"
21     # Open the CSV file in read mode
22     for file in listdir(dir_data):
23         if file.endswith('csv'):
24             with open(file, 'r') as csvfile:
25                 # Read the contents of the CSV file
26                 reader = csv.reader(csvfile)
27                 # Copy all the lines in the file to a list
28                 rows_tempo = list(reader)
29                 for line_1 in rows_tempo[1:]:
30                     in_rows = False
31                     for line_2 in rows :
32                         if line_2[0] == line_1[0] and line_2[1] == line_1[1]
↪ and line_2[2] == line_1[2] and line_2[3] == line_1[3] and line_2[7]
↪ == line_1[7]:
33                             in_rows = True
34                             if int(line_2[5]) > int(line_1[5]):
35                                 line_2[5] = line_1[5]
36                             if int(line_2[6]) > int(line_1[6]):
37                                 line_2[6] = line_1[6]
38                             if int(line_2[7]) > int(line_1[7]):
39                                 line_2[7] = line_1[7]
40                     if not in_rows :
41                         rows.append(line_1)
42
43     with open(name_file+"-final.csv", 'w', newline='') as csvfile:
44         writer = csv.writer(csvfile)
45         writer.writerows(rows)
46
47 def get_auteur(driver, nbr_post):
48     auteur = None
49     data = None
50     try :
51         data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div/div/div/div
↪ /div/div/div/div/div/div[8]/div/div/div[2]/div
↪ /div[2]/div/div[1]").get_attribute("innerHTML")
52     except :

```

```

53         try :
54             data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div/div/div/div
↪ /div/div/div/div/div/div/div[2]/div/div/div[2]/div/div[2]/div/div[1]")
↪ .get_attribute("innerHTML")
55         except:
56             return None
57     try :
58         soup = BeautifulSoup(data, 'html.parser')
59         lst_of_text = soup.findAll(text=True)
60         auteur = lst_of_text[0]
61     except:
62         return None
63     return auteur
64
65
66 def get_is_suggestion(driver, nbr_post):
67     suggestion = False
68     data = None
69     try :
70         data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div/div/div/div
↪ /div/div/div/div/div/div/div[8]/div/div/div[1]/div[7] /div/div/div/span")
71     except:
72         try :
73             data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div/div/div/div
↪ /div/div/div/div/div/div/div[2]/div/div/div[1]/div/div")
74         except:
75             return None
76     try:
77         soup = BeautifulSoup(data, 'html.parser')
78         lst_of_text = soup.findAll(text=True)
79         if lst_of_text[0] == "Suggestions":
80             suggestion = True
81     except:
82         return None
83     return suggestion
84
85

```

```

86 def get_href(driver, nbr_post):
87     data = None
88     try :
89         data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div/div
↪ /div/div/div/div/div/div/div/div[8]/div/div/div[3]")
↪ .get_attribute("innerHTML")
90     except:
91         try :
92
93             data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div
↪ /div/div/div/div/div/div/div/div/div[2]/div/div/div[3]")
↪ .get_attribute("innerHTML")
94         except:
95             pass
96     try :
97         soup = BeautifulSoup(data, 'html.parser')
98         for a in soup.find_all('a', href=True):
99             if a['href'][:24] == "https://www.facebook.com":
100                 return a['href']
101     except:
102         return None
103     return None
104
105 def get_time(driver, nbr_post):
106     Time = None
107     try :
108         data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div/div/div
↪ /div/div/div/div/div/div/div/div[8]/div/div/div[2]
↪ /div/div[2]/div/div[2]/span/span/span[2]/span/a/span/span")
↪ .get_attribute("innerHTML")
109         soup = BeautifulSoup(data, 'html.parser')
110         lst_of_text = soup.findAll(text=True)
111         if lst_of_text[0] == "Sponsoris " :
112             return Time
113         if len(lst_of_text) != 1 :
114             try :
115                 data = driver.find_element(By.XPATH,

```

```

116     ↪ "//body/"+xpath+"/div[" + str(nbr_post) + "]/div/div/div/div/div
117     ↪ /div/div/div/div/div/div/div/div[8]/div/div/div[2]/div/div[2]/div
118     ↪ /div[2]/span/span/span[3]").get_attribute("innerHTML")
119         soup = BeautifulSoup(data, 'html.parser')
120         lst_of_text = soup.findAll(text=True)
121         if lst_of_text[0] == "Sponsoris ":
122             return None
123         return lst_of_text[0]
124     except:
125         pass
126     else:
127         return lst_of_text[0]
128 except:
129     pass
130 try :
131     data = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
132     ↪ str(nbr_post) + "]/div/div/div/div/div
133     ↪ /div/div/div/div/div/div/div/div[2]/div
134     ↪ /div/div[2]/div/div[2]/div/div[2]/span/span/span[2]")
135     ↪ .get_attribute("innerHTML")
136     soup = BeautifulSoup(data, 'html.parser')
137     lst_of_text = soup.findAll(text=True)
138     if lst_of_text[0] == "Sponsoris " :
139         return Time
140     if len(lst_of_text) != 1 :
141         try :
142             data = driver.find_element(By.XPATH,
143             ↪ "//body/"+xpath+"/div[" + str(nbr_post) + "]/div/div/div/div/div/div
144             ↪ /div/div/div/div/div/div/div[2]/div/div/div[2]/div/div[2]/div
145             ↪ /div[2]/span/span/span[3]").get_attribute("innerHTML")
146             soup = BeautifulSoup(data, 'html.parser')
147             lst_of_text = soup.findAll(text=True)
148
149             if lst_of_text[0] == "Sponsoris ":
150                 return None
151             return lst_of_text[0]
152         except:
153             pass
154     else:
155         return lst_of_text[0]

```

165


```

179         data_like = driver.find_element(By.XPATH, "//body/"+xpath+"/div["
↪ + str(nbr_post) + "]/div/div/div/div/div
↪ /div/div/div/div/div/div/div/div/div[8]/div/div/div[4]
↪ /div/div/div/div/div/div[1]/div/div[1]") .get_attribute("innerHTML")
180     except:
181         try :
182             data_like = driver.find_element(By.XPATH,
↪ "//body/"+xpath+"/div[" + str(nbr_post) + "]/div/div/div/div/div
↪ /div/div/div/div/div/div/div/div/div[2]/div/div/div[4]
↪ /div/div/div/div/div/div[1]/div/div[1]") .get_attribute("innerHTML")
183         except:
184             pass
185
186     if data_like != None :
187         soup = BeautifulSoup(data_like, 'html.parser')
188         lst_of_text = soup.findAll(text=True)
189         for elem in lst_of_text:
190             if "," in elem:
191                 elem = elem.replace("\xa0K", "00").replace(",", "")
192             else :
193                 elem = elem.replace("\xa0K", "000")
194             if elem != ' ':
195                 elem = elem.split()
196                 try :
197                     if len(elem) == 1 :
198                         reaction["likes"] = int(elem[0])
199                 except:
200                     pass
201
202     data_com = None
203     try :
204         data_com = driver.find_element(By.XPATH, "//body/"+xpath+"/div[" +
↪ str(nbr_post) + "]/div/div/div/div/div/div/div/div
↪ /div/div/div/div/div/div/div/div[8]/div/div/div[4]/div
↪ /div/div/div/div/div/div[1]/div/div[2]/div[2]") .get_attribute("innerHTML")
205     except:
206         try :
207             data_com = driver.find_element(By.XPATH,
↪ "//body/"+xpath+"/div[" + str(nbr_post) + "]/div/div/div/div/div
↪ /div/div/div/div/div/div/div/div/div[2]/div/div/div[4]/div

```

```

208         ↪ /div/div/div/div[1]/div/div[2]/div[2]") .get_attribute("innerHTML")
209         except:
210             pass
211
212     if data_com != None:
213         soup = BeautifulSoup(data_com, 'html.parser')
214         lst_of_text = soup.findAll(text=True)[0]
215         lst_of_text = lst_of_text.split()
216         if len(lst_of_text) == 2 and lst_of_text[1][0] == "c":
217             reaction["commentaires"] = int(lst_of_text[0])
218
219     data_share = None
220     try :
221         data_share = driver.find_element(By.XPATH, "//body/"+xpath+"/div["
222         ↪ + str(nbr_post) + "]/div/div/div/div/div/div/div
223         ↪ /div/div/div/div/div/div[8]/div/div/div[4]
224         ↪ /div/div/div/div/div[1]/div/div[2]/div[3]")
225         ↪ .get_attribute("innerHTML")
226     except:
227         try :
228             data_share = driver.find_element(By.XPATH,
229             ↪ "//body/"+xpath+"/div[" + str(nbr_post) + "]/div/div/div/div/div
230             ↪ /div/div/div/div/div/div/div[2]/div/div/div[4]
231             ↪ /div/div/div/div/div[1]/div/div[2]/div[3]")
232             ↪ .get_attribute("innerHTML")
233         except:
234             pass
235
236     if data_share != None:
237         soup = BeautifulSoup(data_share, 'html.parser')
238         lst_of_text = soup.findAll(text=True)[0]
239         lst_of_text = lst_of_text.split()
240         if len(lst_of_text) == 2 and (lst_of_text[1][0] == "p" or
241         ↪ lst_of_text[1][0] == "s"):
242             reaction["partages"] = int(lst_of_text[0])
243
244     return reaction
245
246 def main():
247     dir_data = "./"
248     options = webdriver.FirefoxOptions()
249     options.headless = True

```

```

238 driver = webdriver.Firefox(options=options)
239 name_file = None
240 for file in listdir(dir_data):
241     if file.endswith('.html'):
242         nbr_post = 1
243         file_path = 'file://' + os.path.abspath("./" + file)
244         print(file_path)
245         driver.get(file_path)
246         name_file = file.split(".")[0]
247         file_csv = open(name_file+".csv", "w", newline='')
248         writer = csv.writer(file_csv)
249         writer.writerow(["Sponsoris ", "Author", "Time",
↪ "Suggestion", "Description", "Like", "Comment", "Share", "Href"])
250         while True :
251             try :
252                 driver.find_element(By.XPATH, "//body/"+xpath+"/div["
↪ + str(nbr_post) + "]").get_attribute("innerHTML")
253             except:
254                 break
255             auteur = get_auteur(driver, nbr_post)
256             if auteur != None :
257                 auteur = auteur.replace(" & ", "&")
258                 suggestion = get_is_suggestion(driver, nbr_post)
259                 href = get_href(driver, nbr_post)
260                 time = get_time(driver, nbr_post)
261                 sponso = "False"
262                 if time == None:
263                     sponso = "True"
264                 description = get_description(driver, nbr_post)
265                 reactions = get_reactions(driver, nbr_post)
266
267                 writer.writerow([sponso, auteur, time, suggestion,
↪ description, reactions["likes"], reactions["commentaires"],
↪ reactions["partages"], href])
268                 nbr_post += 1
269                 file_csv.close()
270                 creatFinalCsv(name_file)
271 main()

```

A.3. Code for the third experiment

Since we observed that our account was immediately banned, we did not create any specific codes for this experiment. A priori, we would have reused the codes that we had developed for the second experiment A.2.

A.4. Codes for the fourth experiment

This section presents the code that we used to extract information in a .har file.

It is important to note that certain parts of the code are hard-coded as the script was still under development. However, we stopped this experiment for the reasons mentioned in Chapter 2. Therefore, certain parts remained hard-coded and not everything is functioning perfectly.

```
1 import json
2 from haralyzer import HarParser, HarPage
3
4 def get_description(js):
5     try:
6         data_arg = json.dumps(js)
7         data_arg = data_arg.split("\"story\": {\"message\": {\"text\":")
8
9         description = None
10        for i in range(len(data_arg[1])):
11            if data_arg[1][i] == "}":
12                description = data_arg[1][:i]
13                break
14        print(description)
15    except:
16        print("error getting description")
17
18 def get_name(js):
19     try:
20         js = json.dumps(js)
21         js = js.split(",")
22         for i in js:
23             if i[:7] == " \"name\"":
24                 print(i)
```

```

25     except:
26         print("error getting name")
27
28 def get_like(js):
29     try:
30         like = 0
31         js = json.dumps(js)
32         js = js.split(" \\\"i18n_reaction_count\\\":")
33         count = 0
34         for i in range(len(js[1])) :
35             if js[1][i] == "\\\"":
36                 count += 1
37                 if count == 2 :
38                     data = js[1][:i].strip()
39                     print(data[:i].strip("\\\""))
40                     break
41     except:
42         print("error getting like")
43
44 def get_com(js):
45     try:
46         com = 0
47         js = json.dumps(js)
48         js = js.split(" \\\"i18n_comment_count\\\":")
49         count = 0
50         for i in range(len(js[1])) :
51             if js[1][i] == "\\\"":
52                 count += 1
53                 if count == 2 :
54                     data = js[1][:i].strip()
55                     print(data[:i].strip("\\\""))
56                     break
57     except:
58         print("error getting commentaire")
59
60 def get_share(js):
61     try:
62         share = 0
63         js = json.dumps(js)
64         js = js.split(" \\\"i18n_share_count\\\":")

```

```
65     count = 0
66     for i in range(len(js[1])) :
67         if js[1][i] == "\":
68             count += 1
69             if count == 2 :
70                 data = js[1][:i].strip()
71                 print(data[:i].strip("\\"))
72                 break
73     except:
74         print("error getting share")
75
76 def get_time(js):
77     try:
78         js = json.dumps(js)
79         js = js.split("\creation_time\:")
80         if len(js)>1: #else: this means it's sponsored
81             time = js[1][:11]
82             print(time.strip())
83     except:
84         print("error getting time")
85
86 def get_href(js):
87     try:
88         like = 0
89         js = json.dumps(js)
90         js = js.split("\wwwURL\:")
91         count = 0
92         if len(js)>1:
93             for i in range(len(js[1])) :
94                 if js[1][i] == "\":
95                     count += 1
96                     if count == 2 :
97                         data = js[1][:i].strip()
98                         print(data[:i].strip("\\"))
99                         break
100     except:
101         print("error getting href")
102
103 def get_category(js):
104     try:
```

```

105     js = json.dumps(js)
106     js = js.split("\"category\":")
107     count = 0
108     if len(js)>1:
109         for i in range(len(js[1])) :
110             if js[1][i] == "\\":
111                 count += 1
112                 if count == 2 :
113                     data = js[1][:i].strip()
114                     print(data[:i].strip("\\"))
115                     break
116     except:
117         print("error getting category")
118
119 def test_string(js):
120     like = 0
121     js = json.dumps(js)
122     if "Sympathisants" in js:
123         print(True)
124
125 with open('file.har', 'r') as f:
126     har_parser = HarParser(json.loads(f.read()))
127 data = har_parser.har_data
128 data_entries = data["entries"]
129
130 indice = -1
131 var = 1
132 for i in data_entries:
133     indice +=1
134     if i["request"]["url"] == "https://www.facebook.com/api/graphql/":
135         data_posts = i["response"]["content"]["text"]
136         data_posts = data_posts.split("\\r\\n")
137         for js in data_posts:
138             try :
139                 js = json.loads(js)
140                 with open("fichier.json", "w") as file:
141                     json.dump(js, file)
142
143                 get_description(js)
144                 get_name(js)

```

```

145         test_string(js)
146         get_like(js)
147         get_time(js)
148         get_com(js)
149         get_share(js)
150         get_href(js)
151         get_category(js)
152         print("-----")
153         var += 1
154     except:
155         print("error")
156         if var>1:break
157         pass

```

A.5. Codes to create email and Facebook account with a bot

This section contains two subsections. The first subsection includes a script for automatic email creation. The second subsection includes a script that enables the automatic creation of Facebook accounts using pre-existing email and password. These codes work perfectly. You just need to create three files yourself: a file named `Firstname_masculine.txt` with male first names, a file named `Firstname_feminine.txt` with female first names, and a file named `Name.txt` with last names.

A.5.1. Creation of email address

This subsection consists of three files: `Set_server_VPN.py`, `Scrap_creat_account.py` and `main.py`. To enhance our discretion, we aimed to change our IP address (and consequently the country) after creating each email. For this purpose, we developed the script `Set_server_VPN.py`, which utilized commands from NordVPN (the VPN service we employed) to automate the process of changing countries.

Then, we have the script `Scrap_creat_account.py` which utilizes the Selenium module. The

principle behind it is similar to the one explained in Appendix A.1 regarding the first experiment. In this script, we developed a bot using Selenium that navigates on Firefox to create an email on Outlook. The script generates an email address (adapting it if it is already taken), a password, and a birth date. The first name and last name are provided by another script (main.py), which will be explained just after.

Finally, we have the file main.py which randomly selects whether to create a male or female account and based on that, selects a first name and last name from a list (a .txt file) that we created. It then calls the script Scrap_creat_account.py with the name, first name, country, and IP address of the account to be created. If the main.py file is called with the argument -y, it calls the Set_server_VPN.py file to change the location.

A.5.1.1. Set_server_VPN.py

```

1 import os
2 import random
3 import time
4 import socket
5 from selenium.webdriver.common.by import By
6 from selenium import webdriver
7 from selenium.webdriver.firefox.options import Options as FirefoxOptions
8
9 def Select_and_set_server():
10     # Get a list of the possible server of NordVPN
11     lst_servers = []
12     with open("C:/Users/jek/Desktop/TFE_2022-2023/Create_accounts
13     ↪ /Files_txt/Servers.txt") as file:
14         Content = file.read()
15         lst_servers = Content.split("\n")
16
17     # To know wich server (random) to connect
18     random_number = random.randint(0, len(lst_servers)-1)
19     server_to_connect = lst_servers[random_number]
20     os.chdir("C:/Program Files/NordVPN/")
21     os.system('nordvpn -c -g {}'.format(server_to_connect))
22     return server_to_connect

```

```

23 def internet_on(time_start, old_ip_vpn, local_ip, driver, new_country,
    ↪ old_country):
24
25     while time.perf_counter() - time_start < 10 :
26         try:
27             driver.get('https://api.ipify.org/')
28             ip_address = driver.find_element(By.TAG_NAME, "body").text
29
30             if new_country == old_country and ip_address != local_ip:
31                 return ip_address
32             elif ip_address != local_ip and ip_address != old_ip_vpn:
33                 return ip_address
34         except:
35             pass
36     return "False"
37
38 if __name__ == '__main__':
39
40     options = FirefoxOptions()
41     options.add_argument("--headless")
42     driver = webdriver.Firefox(executable_path= r'C:/Users/jek/Desktop
    ↪ /TFE_2022-2023/Create_accounts/OtherFiles/geckodriver.exe',
    ↪ service_log_path="C:/Users/jek/Desktop/TFE_2022-2023
    ↪ /Create_accounts/OtherFiles/geckodriver.log", options=options)
43     driver.get('https://api.ipify.org/')
44     old_ip_vpn = driver.find_element(By.TAG_NAME, "body").text
45
46     hostname = socket.gethostname()
47     local_ip = socket.gethostbyname(hostname)
48
49     file_metadata = open("C:/Users/jek/Desktop/TFE_2022-2023
    ↪ /Create_accounts/Files_txt/Metadatas/metadata.txt", mode="r",
    ↪ encoding="utf-8")
50     lst_info = file_metadata.readlines()
51     file_metadata.close()
52     current_country = lst_info[0].strip()
53
54     new_country = None
55     new_vpn_ip = None
56     connected = False

```

```

57     while not connected :
58         new_country = Select_and_set_server()
59         t = time.perf_counter()
60         new_vpn_ip = internet_on(t, old_ip_vpn, local_ip, driver,
↪ new_country, current_country)
61         if new_vpn_ip != "False":
62             connected = True
63
64         file_new_ip = open("C:/Users/jek/Desktop/TFE_2022-2023
↪ /Create_accounts/Files_txt/Metadatas/metadata.txt", mode="w",
↪ encoding="utf-8")
65         file_new_ip.write(new_country + "\n")
66         file_new_ip.write(new_vpn_ip + "\n")
67         file_new_ip.close()
68         print("Connected to " + str(new_country))
69         print("local IP : " + str(local_ip))
70         print("Old VPN IP : " + str(old_ip_vpn))
71         print("New VPN IP : " + str(new_vpn_ip))

```

A.5.1.2. Scrap_creat_account.py

```

1  import random
2  import string
3  import time
4  import sys
5  from selenium import webdriver
6  from selenium.webdriver.support import expected_conditions as EC
7  from selenium.webdriver.common.by import By
8  from selenium.webdriver.common.keys import Keys
9  from selenium.webdriver.support.wait import WebDriverWait
10 from selenium.webdriver.support.ui import Select
11 from selenium import webdriver
12 from selenium.webdriver.firefox.options import Options as FirefoxOptions
13 from selenium.webdriver.chrome.service import Service as ChromeService
14
15 def generate_email(firstname, name, current_email):
16     email = ""
17     lst_figures= ["0", "1", "2", "3", "4", "5", "6", "7", "8", "9"]
18     if current_email == "":

```

```

19         email += firstname
20         email += name
21         for i in range(4):
22             email += lst_figures[random.randint(0, len(lst_figures)-1)]
23         return email
24
25     current_email += lst_figures[random.randint(0, len(lst_figures)-1)]
26     return current_email
27
28 def generate_password():
29     lst_symboles = []
30     for i in string.printable:
31         lst_symboles.append(i)
32     lst_symboles = lst_symboles[0:94]
33
34     psswd = ""
35     length = random.randint(10, 20)
36     for i in range(length):
37         indice = random.randint(0, len(lst_symboles)-1)
38         psswd += lst_symboles[indice]
39
40     return psswd
41
42 def creat_account(driver, firstname, name):
43     driver.get("https://outlook.live.com/") # we'll be redirected to the
44     ↪ mail service of the country where we are located with the VPN
45     sign_in_button = WebDriverWait(driver,
46     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
47     ↪ "/html/body/header/div/aside/div/nav/ul/li[2]")))
48     sign_in_button.click()
49     creat_on_button = WebDriverWait(driver,
50     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
51     ↪ '//*[@id="signup"]')))
52     creat_on_button.click()
53
54     next_button = WebDriverWait(driver,
55     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
56     ↪ '//*[@id="iSignupAction"]')))
57     email_box = WebDriverWait(driver,
58     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,

```

```

↪ '//*[@id="MemberName"]'))))
51     email = ""
52     error_email = True
53     while error_email:
54         email = generate_email(firstname, name, email)
55         email_box.send_keys(email)
56         next_button.click()
57         time.sleep(1) # TODO : change with a wait page load
58         try :
59             driver.find_element(By.XPATH, '//*[@id="MemberNameError"]')
60             print("Email ERROR")
61         except:
62             error_email = False
63             print("Email ok")
64     print(email)
65
66     next_button = WebDriverWait(driver,
↪ 30).until(EC.element_to_be_clickable((By.XPATH,
↪ '//*[@id="iSignupAction"]'))))
67     password_box = WebDriverWait(driver,
↪ 30).until(EC.element_to_be_clickable((By.XPATH,
↪ '//*[@id="PasswordInput"]'))))
68     error_password = True
69     meta_button = WebDriverWait(driver,
↪ 30).until(EC.element_to_be_clickable((By.XPATH,
↪ '//*[@id="iOptinEmail"]'))))
70     meta_button.click()
71     password = ""
72     while error_password:
73         password = generate_password()
74         password_box.send_keys(password)
75         next_button.click()
76         time.sleep(1) # TODO : change with a wait page load
77         try :
78             driver.find_element(By.XPATH, '//*[@id="PasswordError"]')
79             print("Password ERROR")
80         except:
81             error_password = False
82             print("Password ok")
83     print(password)

```

```

84
85     next_button = WebDriverWait(driver,
86     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
87     ↪ '//*[@id="iSignupAction"]'))))
88     firstname_box = WebDriverWait(driver,
89     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
90     ↪ '//*[@id="FirstName"]'))))
91     name_box = WebDriverWait(driver,
92     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
93     ↪ '//*[@id="LastName"]'))))
94     firstname_box.send_keys(firstname)
95     name_box.send_keys(name)
96     next_button.click()
97
98     next_button = WebDriverWait(driver,
99     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
100    ↪ '//*[@id="iSignupAction"]'))))
101    month_button = WebDriverWait(driver,
102    ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
103    ↪ '//*[@id="BirthMonth"]'))))
104    month_button.click()
105    month = random.randint(2,13)
106    month_choose_button = WebDriverWait(driver,
107    ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
108    ↪ "/html/body/div[1]/div/div
109    ↪ /div[2]/div/div[1]/div[3]/div/div[1]/div[5]/div/div/form/div
110    ↪ /div[4]/div[3]/div[1]/select/option[" + str(month) + "]"))))
111    month_choose_button.click()
112
113    day_button = WebDriverWait(driver,
114    ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
115    ↪ '//*[@id="BirthDay"]'))))
116    day_button.click()
117    day = None
118    if month == 3 :
119        day = random.randint(2,29)
120    else :
121        day = random.randint(2,32)
122    day_choose_button = WebDriverWait(driver,
123    ↪ 30).until(EC.element_to_be_clickable((By.XPATH, "/html/body/div[1]

```

```

106     ↪ /div/div/div[2]/div/div[1]/div[3]/div/div[1]
107     ↪ /div[5]/div/div/div/form/div/div[4]/div[3]/div[2]/select/option[" +
108     ↪ str(day) + "]" )))
109     day_choose_button.click()
110
111     year_box = WebDriverWait(driver,
112     ↪ 30).until(EC.element_to_be_clickable((By.XPATH,
113     ↪ '//*[@id="BirthYear"]'))))
114     year = random.randint(1998, 2004)
115     year_box.send_keys(year)
116
117     next_button.click()
118
119     WebDriverWait(driver,
120     ↪ 1000).until(EC.element_to_be_clickable((By.XPATH,
121     ↪ '//*[@id="KmsiCheckboxField"]'))))
122
123     return [email, password, month-1, day-1, year]
124
125 if __name__ == '__main__':
126
127     # Get metadata
128     firstname = sys.argv[1].replace(" ", "")
129     name = sys.argv[2].replace(" ", "")
130     country = sys.argv[3].replace(" ", "") # remove eventual space. Ex :
131     ↪ San Francisco --> SanFrancisco
132     ip = sys.argv[4].replace(" ", "")
133     male_fem = sys.argv[5]
134     if str(male_fem) == "0":
135         male_fem = "F"
136     else :
137         male_fem = "M"
138
139     # Creat driver
140     options = FirefoxOptions()
141     options.set_preference('intl.accept_languages', 'en-US')
142     options.add_argument("--disable-gpu")
143     options.add_argument("--start-fullscreen")
144     driver = webdriver.Firefox(executable_path=
145     ↪ r'C:/Users/jek/Desktop/TFE_2022-2023

```

```

138     ↪ /Create_accounts/OtherFiles/geckodriver.exe',
139     ↪ service_log_path="C:/Users/jek/Desktop/TFE_2022-2023/Create_accounts
140     ↪ /OtherFiles/geckodriver.log", options=options)
141     driver.delete_all_cookies()
142     email, password, month, day, year = creat_account(driver, firstname,
143     ↪ name)
144     print(firstname)
145     print(name)
146     print(country)
147     print(ip)
148     print(email)
149     print(password)
150     print(month)
151     print(day)
152     print(year)
153     s = "\n" + str(male_fem) + " " + str(firstname) + " " + str(name) + "
154     ↪ " + str(country) + " " + str(ip) + " " + str(email) + "@outlook.com
155     ↪ " + str(password) + " " + str(month) + " " + str(day) + " " +
156     ↪ str(year)
157     file_account =
158     ↪ open("C:/Users/jek/Desktop/TFE_2022-2023/Create_accounts
159     ↪ /Files_txt/Accounts.txt", mode="a", encoding="utf-8")
160     file_account.write(s)
161     file_account.close()
162     driver.close()

```

A.5.1.3. main.py

```
1 import os
2 import random
3 import time
4 import socket
5 import sys
6
7 def change_ip():
8     os.system('python "./Set_server_vpn.py"')
9
10 def get_info_ip():
```



```

11     file_new_ip = open("C:/Users/jek/Desktop/TFE_2022-2023
    ↪ /Create_accounts/Files_txt/Metadatas/metadata.txt", mode="r",
    ↪ encoding="utf-8")
12     lst_info = file_new_ip.readlines()
13     file_new_ip.close()
14     return lst_info
15
16 def creat_account(firstname, name, country, ip):
17     os.system('python ./Scrap_creat_account.py {} {} {}
    ↪ {}'.format(firstname, name, country, ip))
18
19 def get_firstname_name():
20     # Get the account already created
21     file_accounts =
    ↪ open("C:/Users/jek/Desktop/TFE_2022-2023/Create_accounts
    ↪ /Files_txt/Accounts.txt", mode="r", encoding="utf-8")
22     lines_accounts = file_accounts.readlines()
23     file_accounts.close()
24
25     # Get the firstname
26     Is_men_women = random.randint(0, 1) # 0 = we will create a men
    ↪ account. 1 = we will create a women account.
27
28     file_firstname = None
29     if Is_men_women == 0 :
30         file_firstname =
    ↪ open("C:/Users/jek/Desktop/TFE_2022-2023/Create_accounts
    ↪ /Files_txt/Firstname_masculine.txt", mode="r", encoding="utf-8")
31     else :
32         file_firstname = open("C:/Users/jek/Desktop/TFE_2022-2023
    ↪ /Create_accounts/Files_txt/Firstname_feminine.txt", mode="r",
    ↪ encoding="utf-8")
33     lines_firstname = file_firstname.readlines()
34     file_firstname.close()
35
36     # Get the name
37     file_name = open("C:/Users/jek/Desktop/TFE_2022-2023/Create_accounts
    ↪ /Files_txt/Name.txt", mode="r", encoding="utf-8")
38     lines_name = file_name.readlines()
39     file_name.close()

```

```
40
41     not_unique = True
42     while not_unique:
43         not_unique = False
44         random_firstname = random.randint(0, len(lines_firstname)-1)
45         random_name = random.randint(0, len(lines_name)-1)
46
47         firstname = lines_firstname[random_firstname].strip()
48         name = lines_name[random_name].strip()
49
50         for line in lines_accounts:
51             line = line.split(',')
52             # If it exist already an account with this firstname and name
53             if firstname == lines_accounts[0].strip() and name ==
↪ lines_accounts[1].strip() :
54                 not_unique = True
55                 break
56
57         return [firstname, name]
58
59 def get_new_user():
60     firstname, name = get_firstname_name()
61     return [firstname, name]
62
63 if __name__ == '__main__':
64     if sys.argv[1] == "-y":
65         change_ip()
66
67     lst_info = get_info_ip()
68     country = lst_info[0].strip()
69     ip = lst_info[1].strip()
70
71     firstname, name = get_new_user()
72
73     creat_account(firstname, name, country, ip)
```

A.5.2. Creation of Facebook account

This subsection contains a single file that creates a bot using the Selenium module. The bot is designed to create a Facebook account using pre-created email and password.

A.5.2.1. Creat_Facebook_account.py

```
1 import random
2 import string
3 import time
4 import sys
5 import os
6 from selenium import webdriver
7 from selenium.webdriver.support import expected_conditions as EC
8 from selenium.webdriver.common.by import By
9 from selenium.webdriver.common.keys import Keys
10 from selenium.webdriver.support.wait import WebDriverWait
11 from selenium.webdriver.support.ui import Select
12 from selenium import webdriver
13 from selenium.webdriver.firefox.options import Options as FirefoxOptions
14 from selenium.webdriver.chrome.service import Service as ChromeService
15
16 def change_city(city):
17     if city == "NewZealand":
18         city = "New Zealand"
19     if city == "NorthMacedonia":
20         city = "North Macedonia"
21     if city == "SanFrancisco":
22         city = "San Francisco"
23     if city == "BosniaandHerzegovina":
24         city = "Bosnia and Herzegovina"
25     if city == "CostaRica":
26         city = "Costa Rica"
27     if city == "CzechRepublic":
28         city = "Czech Republic"
29     if city == "HongKong":
30         city = "Hong Kong"
31     if city == "SouthAfrica":
32         city = "South Africa"
```

```

33     if city == "SouthKorea":
34         city = "South Korea"
35     if city == "UnitedArabEmirates":
36         city = "United Arab Emirates"
37     if city == "UnitedKingdom":
38         city = "United Kingdom"
39     if city == "UnitedStates":
40         city = "United States"
41     if city == "KansasCity":
42         city = "Kansas City"
43     if city == "LosAngeles":
44         city = "Los Angeles"
45     if city == "NewYork":
46         city = "New York"
47     if city == "SaintLouis":
48         city = "Saint Louis"
49     if city == "SaltLakeCity":
50         city = "Salt Lake City"
51     return city
52
53 def change_ip(city, driver):
54     city = change_city(city)
55     os.chdir("C:/Program Files/NordVPN/")
56     os.system('nordvpn -c -g {}'.format(city))
57     os.chdir("C:/Users/jek/Desktop
↪ /TFE_2022-2023/Create_accounts/Creat_Facebook_account")
58     time.sleep(2)
59     not_connected = True
60     while not_connected :
61         try:
62             driver.get('https://api.ipify.org/')
63             driver.find_element(By.TAG_NAME, "body").text
64             not_connected = False
65         except:
66             time.sleep(1)
67         pass
68     driver.delete_all_cookies()
69
70 def creat_account(male_fem, firstname, name, email, password, day, month,
↪ year, driver):

```

```

driver.get("https://www.facebook.com/")

try:
    cookies_button = WebDriverWait(driver,
    ↪ 10).until(EC.element_to_be_clickable((By.XPATH,
    ↪ '/html/body/div[3]/div[2]/div/div/div/div/div[3]/button')))
    cookies_button.click()
except :
    pass

creat_account_button = WebDriverWait(driver,
    ↪ 30).until(EC.element_to_be_clickable((By.XPATH, '/html/body/div[1]
    ↪ /div[1]/div[1]/div/div/div/div[2]/div/div[1]/form/div[5]/a')))
creat_account_button.click()

firstname_box = WebDriverWait(driver,
    ↪ 30).until(EC.element_to_be_clickable((By.XPATH, '/html/body/div[3]
    ↪ /div[2]/div/div/div[2]/div/div/div[1]/form
    ↪ /div[1]/div[1]/div[1]/div[1]/div/input'))))
firstname_box.send_keys(firstname)

name_box = WebDriverWait(driver,
    ↪ 30).until(EC.element_to_be_clickable((By.XPATH, '/html/body/div[3]
    ↪ /div[2]/div/div/div[2]/div/div/div[1]/form
    ↪ /div[1]/div[1]/div[1]/div[2]/div/div[1]/input'))))
name_box.send_keys(name)

email_box = WebDriverWait(driver,
    ↪ 30).until(EC.element_to_be_clickable((By.XPATH, '/html/body/div[3]
    ↪ /div[2]/div/div/div[2]/div/div/div[1]/form
    ↪ /div[1]/div[2]/div/div[1]/input'))))
email_box.send_keys(email)

re_enter_email_box = WebDriverWait(driver,
    ↪ 30).until(EC.element_to_be_clickable((By.XPATH, '/html/body
    ↪ /div[3]/div[2]/div/div/div[2]/div/div/div[1]/form
    ↪ /div[1]/div[3]/div/div/div[1]/input'))))
re_enter_email_box.send_keys(email)

password_box = WebDriverWait(driver,
    ↪ 30).until(EC.element_to_be_clickable((By.XPATH, '/html/body/div[3]

```



```
147     ↪ /Create_accounts/OtherFiles/geckodriver.exe',  
148     ↪ service_log_path="C:/Users/jek/Desktop/TFE_2022-2023/Create_accounts  
149     ↪ /OtherFiles/geckodriver.log", options=options)  
150         driver.delete_all_cookies()  
151  
152         change_ip(city, driver)  
  
153         creat_account(male_fem, firstname, name, email, password, day,  
154         ↪ month, year, driver)  
155         break
```

A.5.3. Screenshots of banned accounts

As a reminder, the emails were automatically created using a Python script that utilized the Selenium module. The only manual step involved was solving the captchas manually.

A.5.3.1. Facebook accounts created with a bot

In the screenshots, you may notice that the search bar appears in red. This indicates that Firefox itself detected that it was being controlled by an automated script.

Figure A.1.: First attempt.

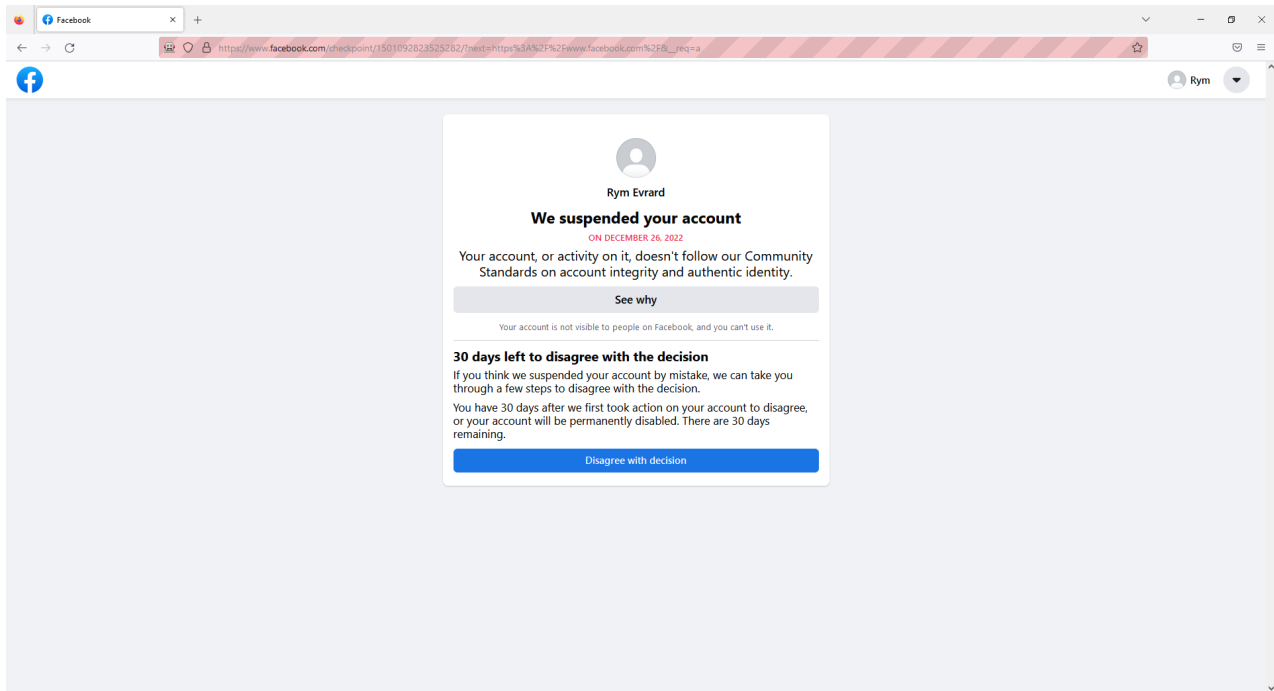
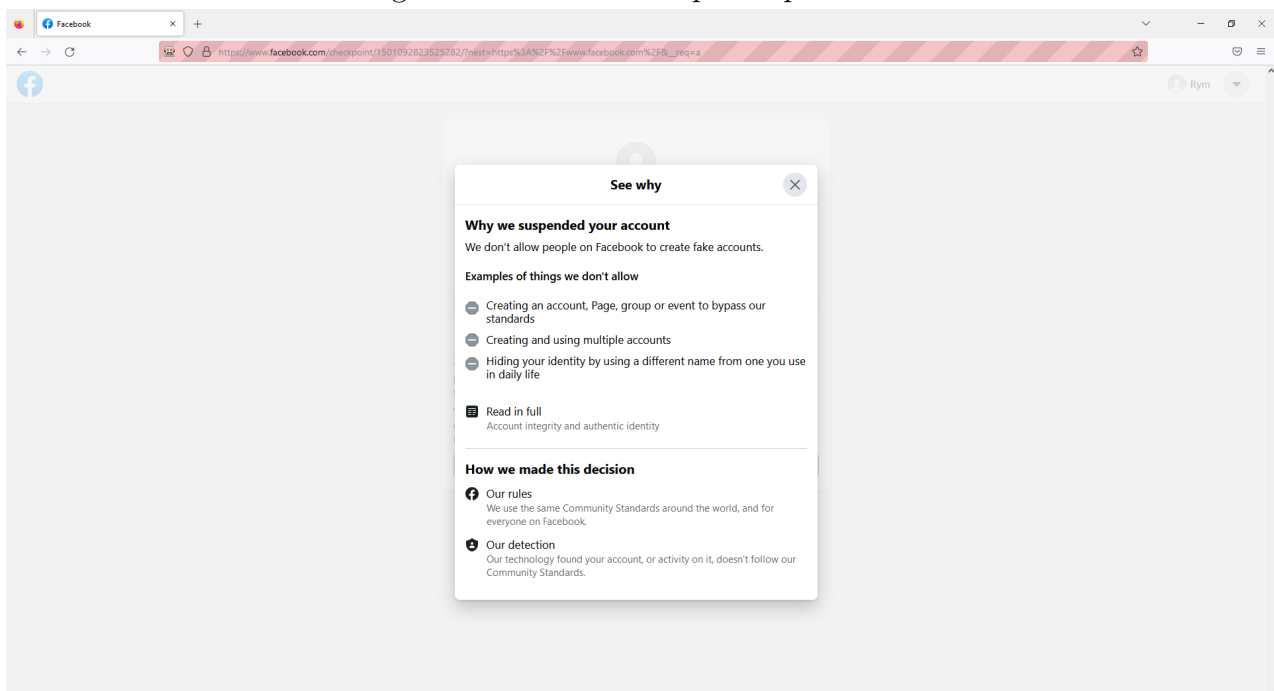


Figure A.2.: First attempt - explanation.



A.5.3.1.1. First attempt

Figure A.3.: Second attempt.

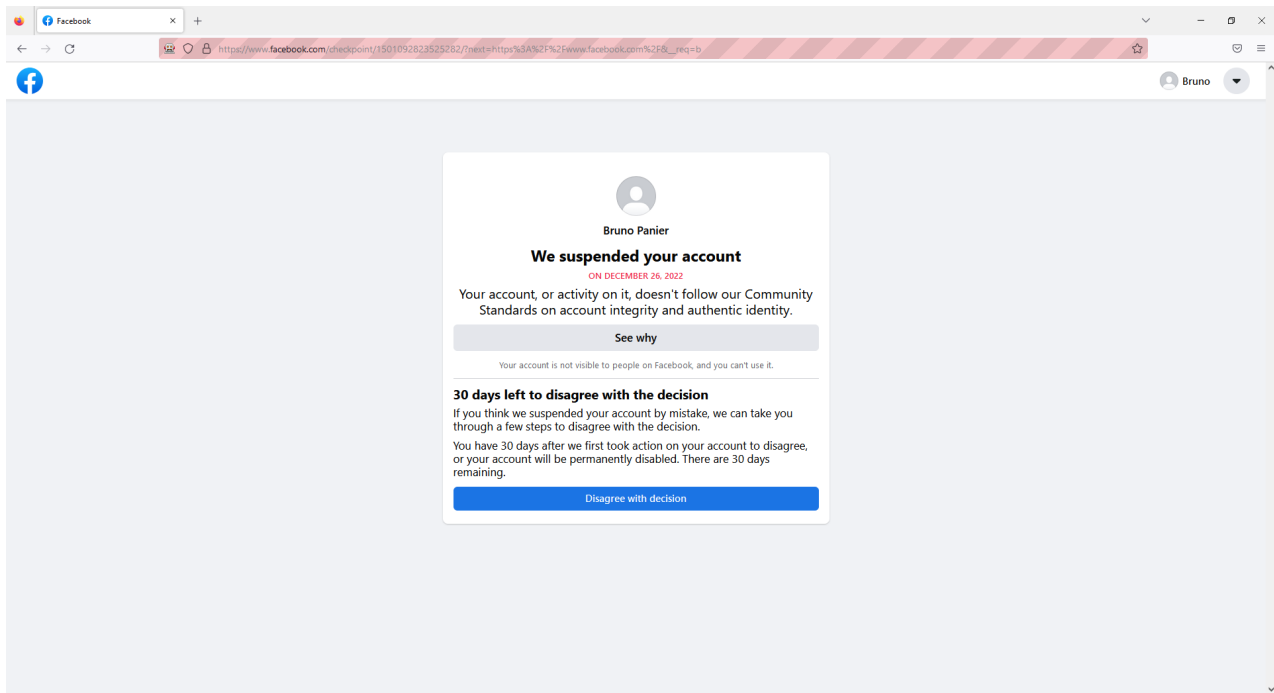
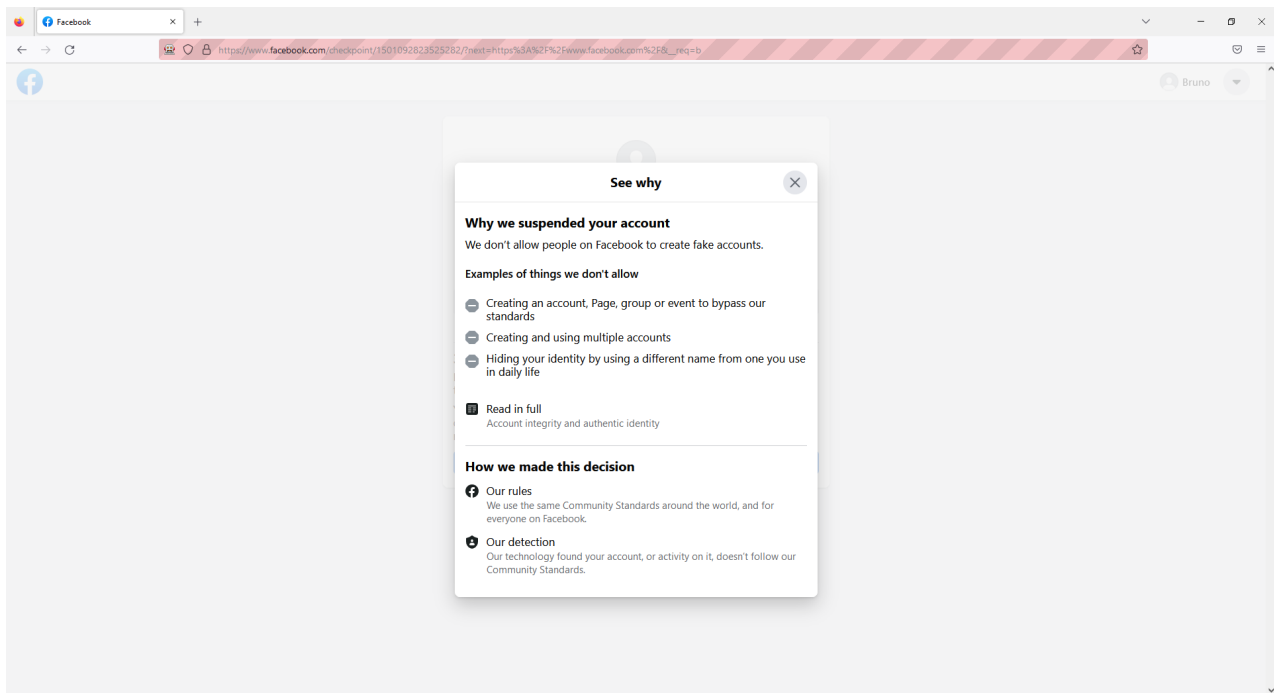


Figure A.4.: Second attempt - explanation.



A.5.3.1.2. Second attempt

A.5.3.2. Facebook accounts created manually

Here, we used an email created with a bot, but we attempted to create a Facebook account manually.

You may notice that the search bar is no longer red. This proves that we created it manually.

A screenshot of a web browser displaying a Facebook account suspension notification. The browser's address bar shows a URL from facebook.com. The page features a white notification card on a light blue background. At the top of the card is a grey profile picture placeholder and the name 'Rym Evrard'. Below this, the heading 'We suspended your account' is followed by the date 'ON DECEMBER 26, 2022'. The main text states: 'Your account, or activity on it, doesn't follow our Community Standards on account integrity and authentic identity.' A grey button labeled 'See why' is positioned below the text. A smaller line of text reads: 'Your account is not visible to people on Facebook, and you can't use it.' Further down, the section '30 days left to disagree with the decision' is followed by explanatory text: 'If you think we suspended your account by mistake, we can take you through a few steps to disagree with the decision. You have 30 days after we first took action on your account to disagree, or your account will be permanently disabled. There are 30 days remaining.' At the bottom of the card is a prominent blue button labeled 'Disagree with decision'.

A screenshot of a web browser displaying a Facebook account suspension notification. The browser's address bar shows a URL from facebook.com. The notification is a white modal dialog box with a close button (X) in the top right corner. The title of the dialog is "See why". The main heading inside the dialog is "Why we suspended your account", followed by the text "We don't allow people on Facebook to create fake accounts." Below this, there is a section titled "Examples of things we don't allow" which contains a list of three items, each preceded by a minus sign icon: "Creating an account, Page, group or event to bypass our standards", "Creating and using multiple accounts", and "Hiding your identity by using a different name from one you use in daily life". Underneath the list is a link that says "Read in full" with the subtext "Account integrity and authentic identity". The next section is titled "How we made this decision" and contains two items, each preceded by a Facebook 'f' logo icon: "Our rules" with the text "We use the same Community Standards around the world, and for everyone on Facebook.", and "Our detection" with the text "Our technology found your account, or activity on it, doesn't follow our Community Standards." The background of the browser shows the Facebook homepage with a search bar and a navigation menu.

A.5.3.3. Email and Facebook accounts created manually

Even when we attempted to create an account manually, we were still banned.

Figure A.7.: Account banned with email and Facebook account created manually

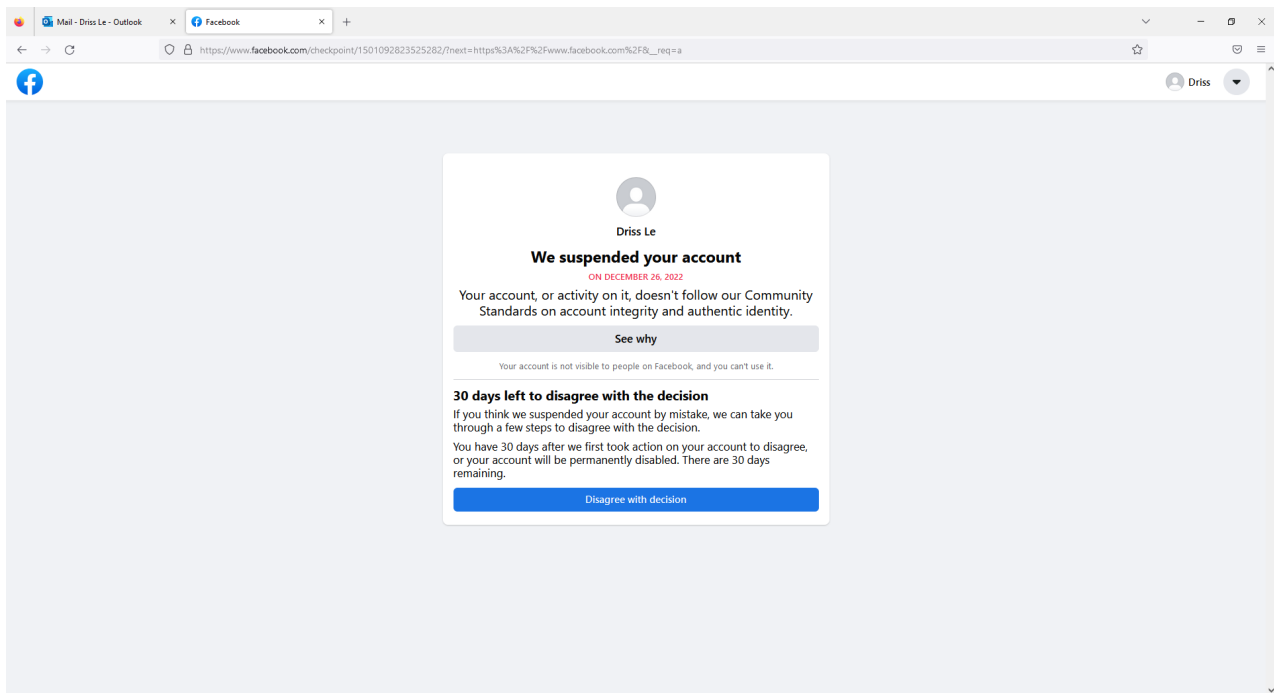
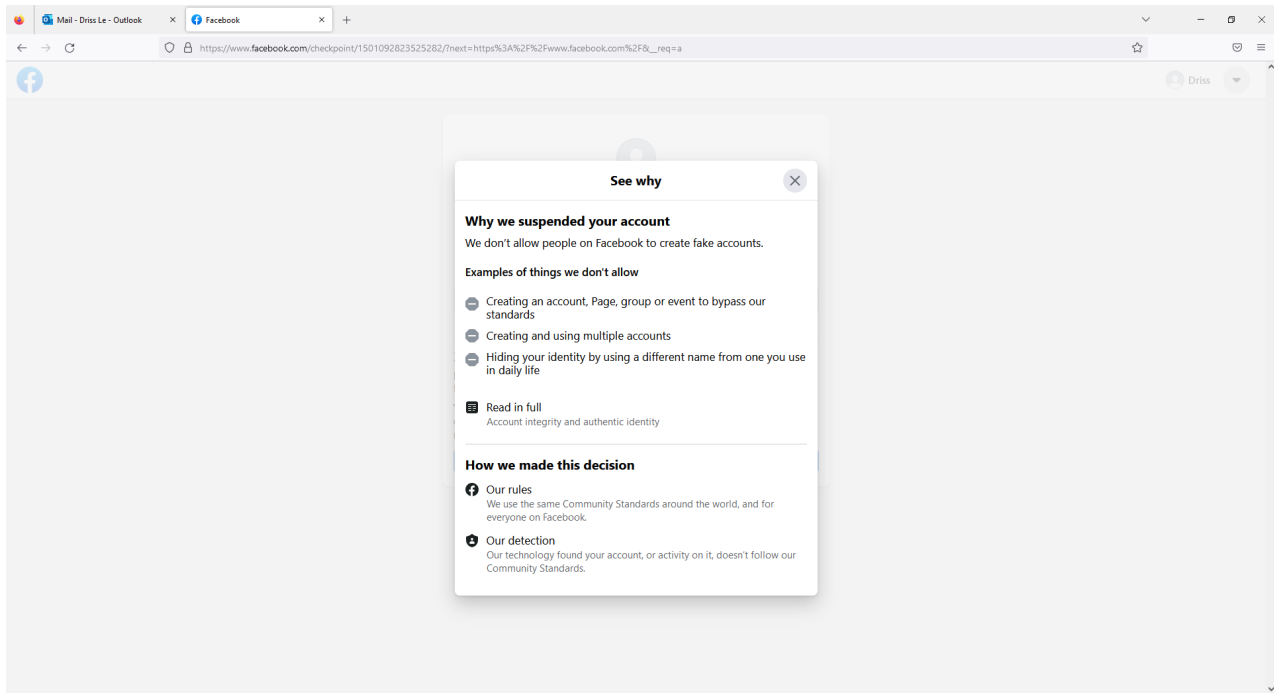


Figure A.8.: Account banned with email and Facebook account created manually - explanation



B. Case studies

B.1. Evolution of our society with the arrival of the Internet and new technologies

B.1.1. The incredible evolution of technologies

In the last 30 years, technology has continued to evolve in an exponential way.

The Internet was created in 1969 in order to link computers together and allow the exchange of information through protocols and addresses [242]. It was in 1989 that Tim Berners-Lee created the first version of the Web, which allows users to consult pages via a browser [309][253][152].

This arrival of the Internet and the web has revolutionized the world. First, we can see this by looking at the evolution of the number of users over time. As shown on the graph show in subsection B.1.3, in 1991, the world population was around 5.40 billion, of which 26 million were using the Internet (which is equivalent to a little less than 0.5% of the population). In 2021, there were about 7.91 billion people, of which about 4.7 billion were using the Internet (i.e. a little more than 63% of the population). However, if we look at figure B.3, we notice that it is in the most developed countries where the Internet has developed the most and has become an essential part of people's lives. For example, it took only 10 years for countries like Canada and Norway to have more than 50% of their population using the Internet. In 2016, many countries such as France or the United Kingdom had more than 80% of the population using the Internet. Internet has quickly become a part of the customs of several countries.

Now, if we take a closer look at the evolution of technology, we also notice a phenomenal growth that human beings have never had before. The processing power of computers and

the number of transistors in micro-processors have almost doubled every 2-3 years as shown in the figures B.7 [19]. To give some representative figures and according to the graph in the subsubsection B.1.3.1: in 1971, the number of transistors in a chip was around 2300 and in 2017 it has increased to 19.7 billion. In terms of computer power, we have gone from 124 gigaFlops in 1993 to 1.10 billion gigaFlops in 2022 [19]. If we now look at memory, we see the same thing: the cost of a megabyte of memory has dropped from \$100 million to less than \$0.01 in 2008 as shown in figure B.7. This decrease in the price of memory and the increase in computing power have allowed many companies such as Google, Meta, Microsoft, etc. to create huge data centers where a lot of data is collected and analyzed. In 2021, Meta has 18 data centers in the world with a total area of about 40 million square feet [247]. In 2022, this number of servers allows Meta to generate 4 petabytes (or 4000 terabytes) of data every day [217][75].

Other interesting figures are the amount of data that have been generated each year on the web. The trend is the same: our production doubles almost every 2-3 years (as it's shown on figure B.9) and we have reached 94 zettabytes of data produced during the year 2022 [278] compared to 44 zettabytes in 2020 [195].

Another factor is the arrival of smartphones. Indeed, smartphones have also revolutionized the way we consume the Internet. The first smartphone that is similar to those we have today is the iPhone which appeared in 2007 [173]. You know the rest: in France, nearly 90% of the population has a smartphone [35] as shown in figure B.10. It's of course the richest countries that use the most smartphones, while the poorest countries can go down to a percentage of 50% [29]. In 2022, according to the ONU, almost three quarters of children over 10 years old will have a smartphone and therefore access to the Internet[29]. We can see on the figure B.10 that in 2020, almost 92% of the French aged 12 to 17 years have a smartphone [35]. Smartphones have made it possible to access the Internet and various websites wherever you are via a cellular connection. We had a democratization of the use of the Internet outside a fixed place like an office or at home where we had to use a computer to access this service. Indeed, we can see it when we walk in the streets, we go to a coffee, take the subway, ... a large majority of people have their eyes riveted on their smartphone. As a personal observation, when I take the subway in Brussels, I see that people are mostly on social networks such as Facebook, TikTok or Instagram. If we look at the numbers, it confirms the fact that smartphones have taken an important place in our lives. Indeed, around 65% of the Internet

traffic comes from smartphones against around 32% that come from computers (the rest being tablets, ...) in 2023 [34].

In 2023, an American spends an average of 5.4 hours a day on his smartphone [129]. In 2021, in the world, the population was spending around 4 hours per day on their smartphone [73][135] and in the USA, for people over 18 years old, a majority spends between 5 and 6 hours a day on their screen as shown in figures in subsection B.1.6.

Now, it would be interesting to look at the question of who controls what in this network? Indeed, we consume mainly via the Internet but what do we consume? Is there a diversification or a monopolization of the different actors present on the Internet? Internet being a major actor in our society and communication, it is an important question that we will answer now. We'll answer those questions: first, we'll look at what major browsers people are using. Indeed, they are responsible for ranking and selecting which websites to highlight. Then, what are the most visited websites? and are there actors responsible for a large majority of the traffic on the network? Is there a diversification of those actor? Also, what are the most used smartphone applications? In fact, as said before, we spend a lot of time on these devices.

Without surprise, the most used browser is Google Chrome with around 60-65% of the market in 2022 and followed by Safari with 15-18% depending on the period [31][32]. If we focus on smartphones, 96% of the traffic concerning Internet searches are done via Google in 2022 in France [240].

Next, let's look at the most visited website¹, again and without surprise, in 2022, we have giants such as Google, Facebook (i.e. Meta), Youtube (which belongs to Google), Twitter, Instagram (i.e. Meta) or Reddit which are in the top of the consulted websites [112][11][112] as shown in figures in subsection B.1.7.

Let's now focus on the smartphone. What are the most downloaded applications? From 2020 to 2022, we find mainly TikTok, Instagram (i.e. Meta), Facebook (i.e. Meta), WhatsApp (i.e. Meta), Messenger (i.e. Meta), SnapChat and Telegram [211][212][210].

What about Internet traffic via smartphones? In 2021, we see that 20.4% of the traffic comes from Youtube (i.e. Google) and 5.4% directly from Google, 17.5% from Facebook and Facebook video (i.e Meta), 6.8% from TikTok, 5.1 from Instagram (i.e Meta), 1.8% from WhatsApp

¹Please note that we are not talking about traffic on the network but **visits** of web pages. In the different studies found, generally they do not take into account messengers such as Messenger or WhatsApp which are more accessed via applications than websites (same for Netflix) and we will come back to this later.

(i.e Meta) and 1.8% from Netflix [340]. In the world, videos represent 43% of the traffic with Youtube and Netflix as leaders. Then we have the social networks that occupy 12.69% of traffic with Meta and TikTok as the main leader. The web searches and gaming take 15% of the traffic with Google as the main browser for searches. Messaging occupies 5% of the traffic with Meta (Messenger and WhatsApp) as the leader[33]. So we have almost 55% of the Internet traffic coming from 6 main players: Google, Netflix, Facebook (i.e. Meta), Apple, Amazon and Microsoft (in other words: GAFAM + Netflix)[33]. If we look more precisely in France: Netflix, Google, Amazon and Meta occupy nearly 50% of the Internet traffic² [104]

So, on one hand, our society has completely changed since the arrival of the Internet and all these new technologies. We don't consume the same way product and services as in the 80's/90's or 00's, we spend hours and hours on our screens and our main communication is through our smartphone. We only have to walk in the subway to see that at least 70% of people have their eyes on their mobile phone and scroll on their favorite social network. We are now in a world where communication and consumption of services is mainly done through our devices rather than through real social links. We all have in our pocket this object that allows us, when we are bored, to take it and spend our time consuming what is offered to us for free.

On the other hand, we notice that all of what we consume is mainly controlled by the GAFAM + Netflix. With the increasing power of computers, the very low cost of memory and the creation of the Internet network allowing the circulation of data, these companies have managed to monopolize this sector. They control the majority of the information flow, control what is offered to us and collect all our data on how we use their services. Wouldn't there be some manipulation in what we are offered? Moreover, they are equipped with a big power of calculation and AI is in full rise. For example, between 2014 and 2016, Google bought 9 AI companies, making it one of the pioneers in this field [357]. The five largest technology companies have enough capital to crowd out their competitors: start-ups, universities, municipalities, companies in other industries and poorer countries [357]. This gives them even more power to control the flow of information.

²Here is a link to an interesting 87-page report showing the state of the Internet in France in 2017 : https://www.arcep.fr/uploads/tx_gspublication/rapport-etat-internet-france-2017-mai2017.pdf

B.1.2. Arrival of social networks

According to some, the first social network arrived in 1997: Six Degrees [189][66]. But it is only from 2004 that we start to see social networks such as Facebook and Youtube appear, which became the two main actors in the sphere of social networks [154].

In this thesis, we will focus on Facebook (i.e. Meta). Why do we do this? Because it is the leading and most controversial social network as we will see later. As shown in the figure in subsection B.1.8, Facebook is the leading social network in terms of monthly active users. In July 2022, Facebook reached almost 3 billion active users worldwide [111] (be careful, all the figures that are given are to be questioned because it certainly includes fake accounts, bots, ... that Facebook has not detected). So let's keep a critical mind. The figures of this year are interesting to know but what is more judicious to note is that it is the leader since 2010 with Youtube that follows closely. To this we must add the fact that they bought Instagram in 2012 [59] and WhatsApp in 2014 [285] and own Messenger. This gives them a certain monopoly in this market and allows them to have a good understanding of how the flow of information is circulating and the intensity of the links between all its users. Moreover, it is indeed the social network where people spend the most time (on average 33 min in 2023 [130]). This is why in this thesis we focus on Facebook and more specifically on our experiments that we have explained to you in order to understand the impact that social networks have on our society. This will not prevent us from mentioning other social networks such as Twitter in the analysis of previous articles on the subject. Interestingly, Facebook is declining among younger people (<25) who are going more on Instagram and TikTok.³

Other interesting graphics can be found at the following source: [282].

³For your information: Facebook was created in 2004 but was limited to Harvard students and it is in 2006 that Facebook becomes accessible to all [209]. It is around 2008-2010 when social networks began to develop on mobile applications [106][179].

B.1.3. Evolution of the world population and web users

Figure B.1.: Evolution of the population from 1990 to 2021 [48].

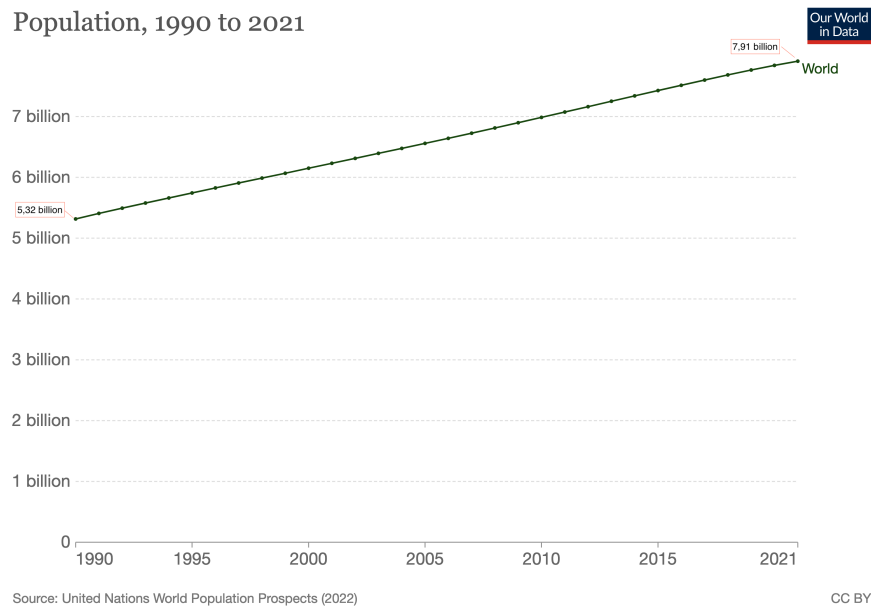


Figure B.2.: Evolution of the number of user of Internet from 1991 to 2022 [324].

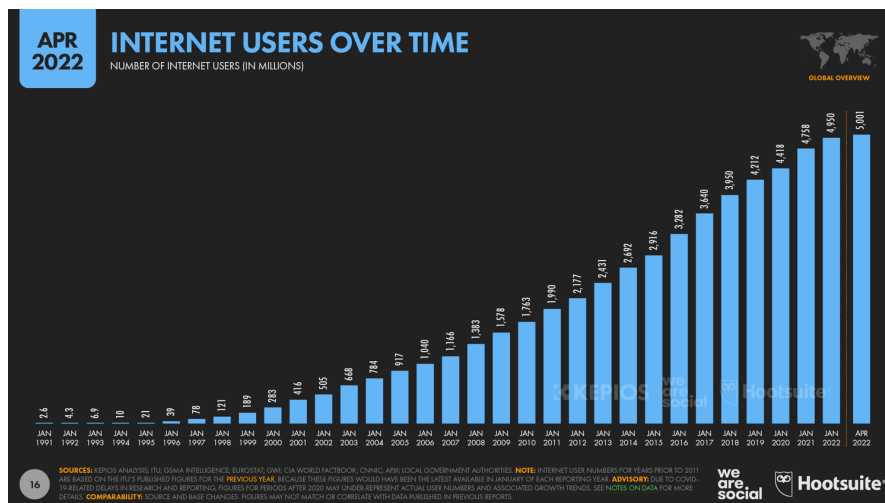
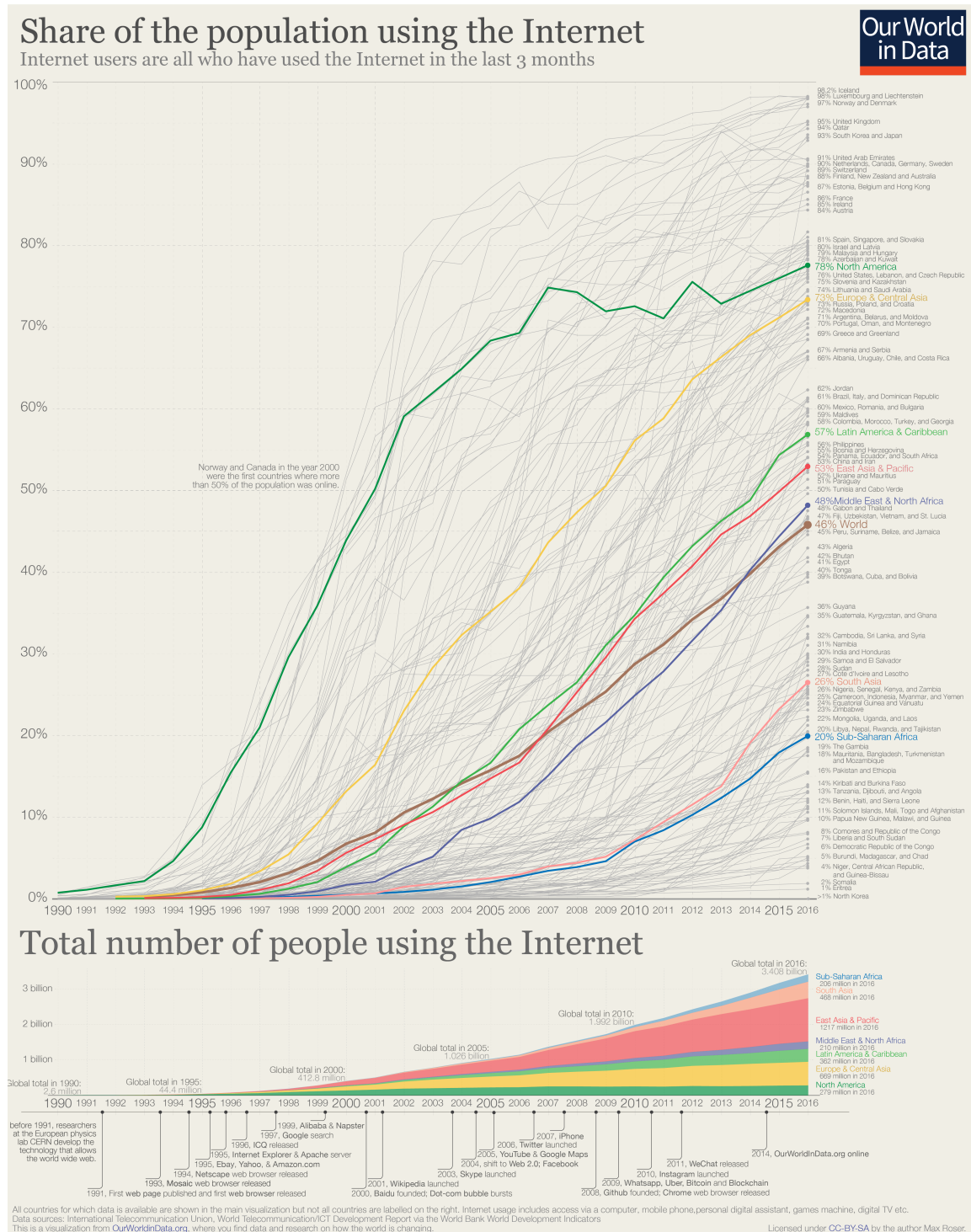


Figure B.3.: Share of Internet users [253].



B.1.3.1. The growth of computer processing power and memory cost

Figure B.4.: Transistors per microprocessor (logarithmic scale) [46].

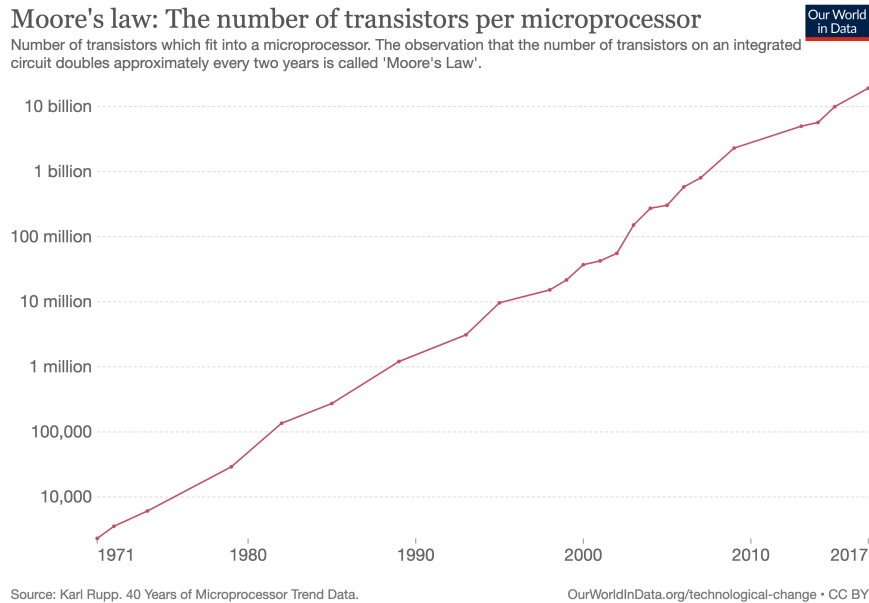


Figure B.5.: Transistor count over time (logarithmic scale) [45].

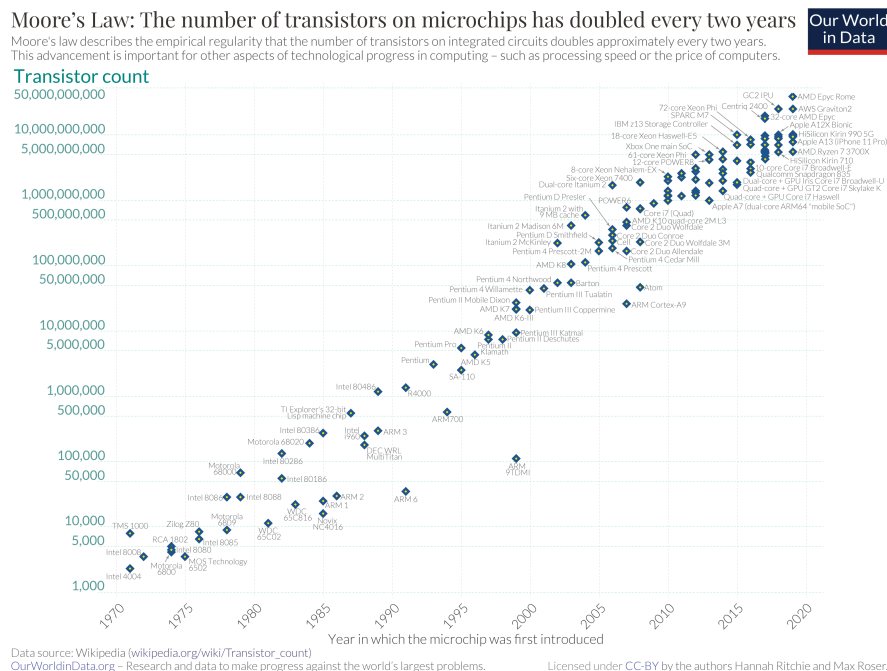


Figure B.6.: NVIDIA - computer processing power (linear scale) [264].

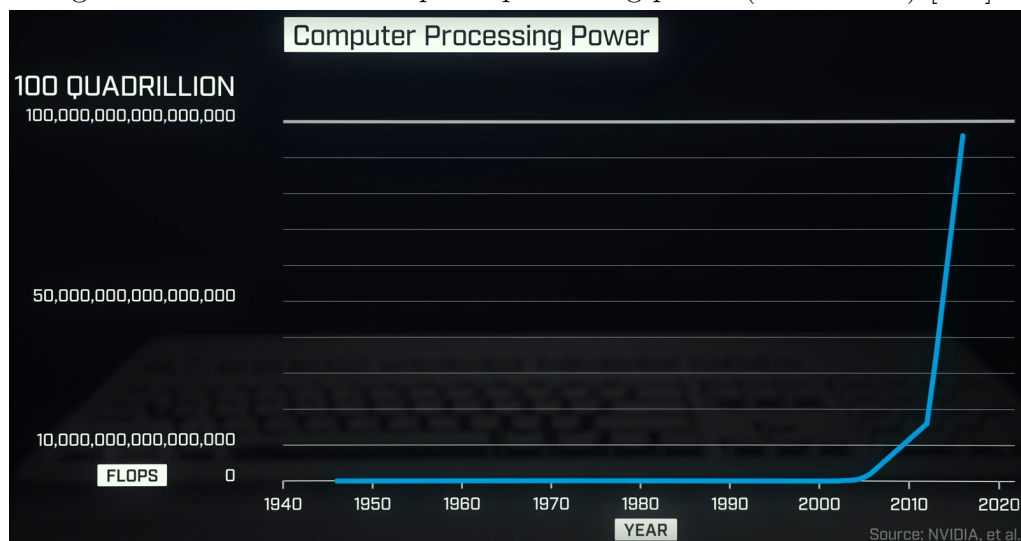


Figure B.7.: Supercomputer power flops over time (logarithmic scale). Y-Axis express in gi-gaFlops [19].

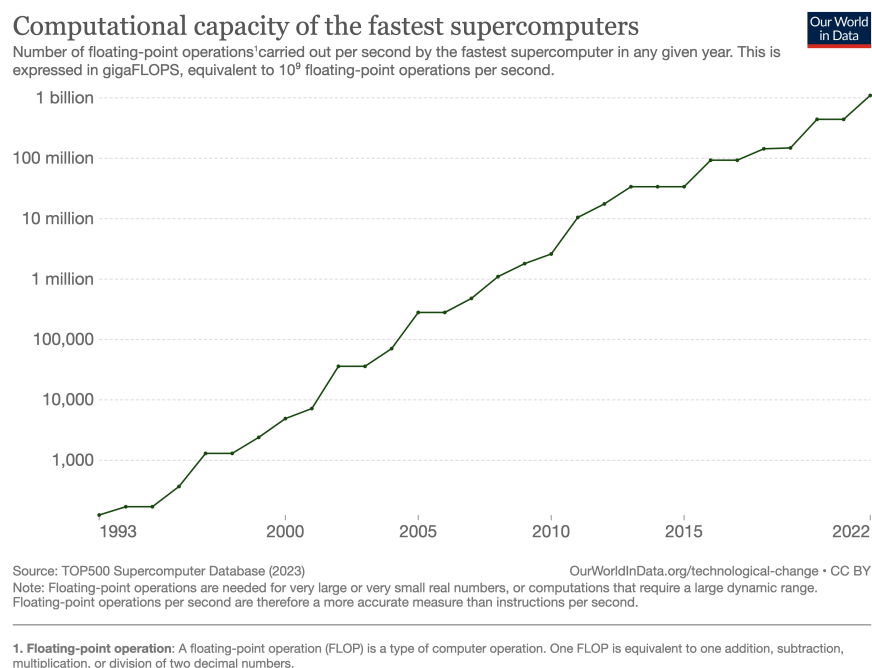
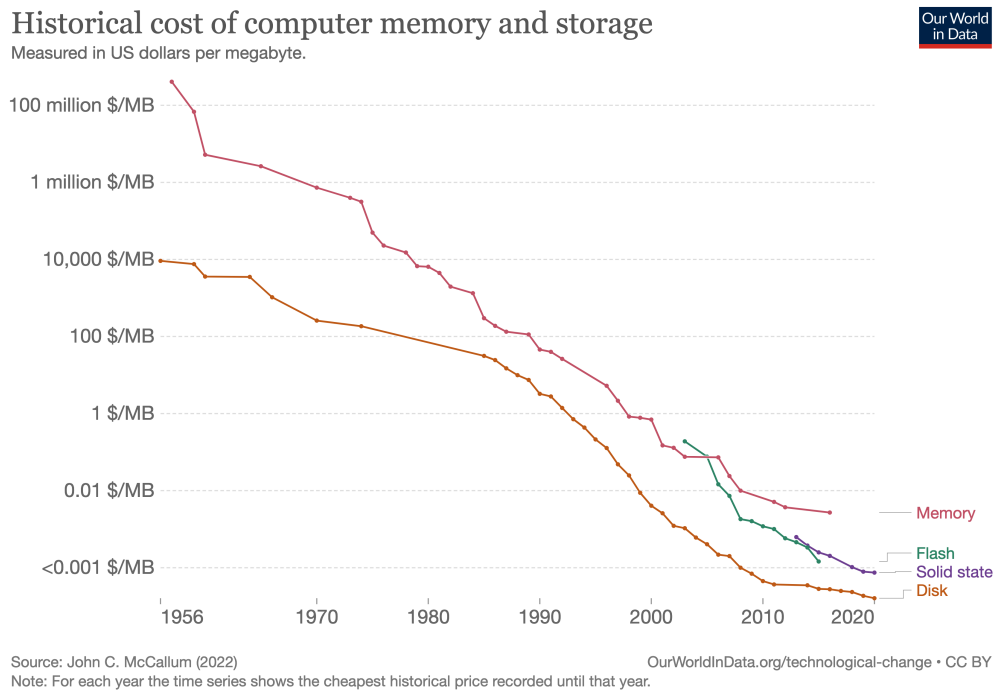
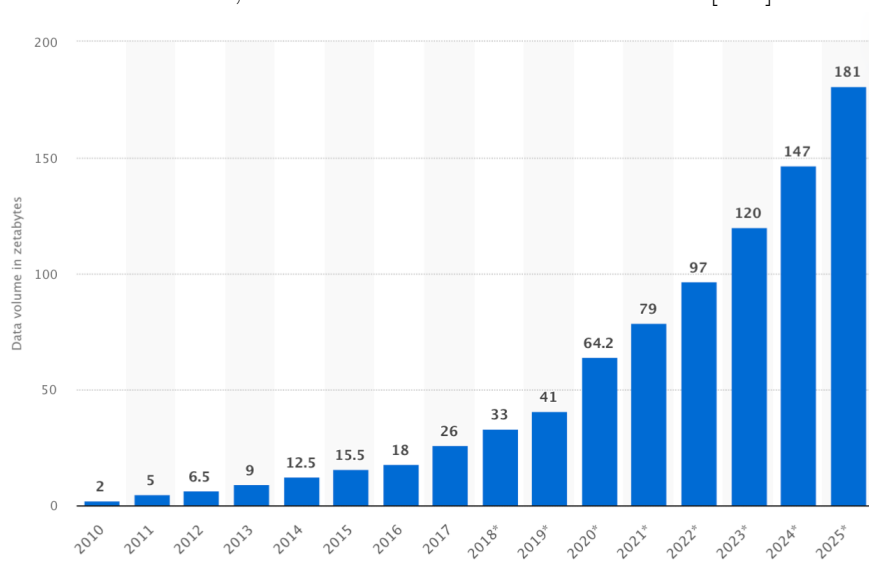


Figure B.8.: Historical cost of computer memory and storage (logarithmic scale) [17].



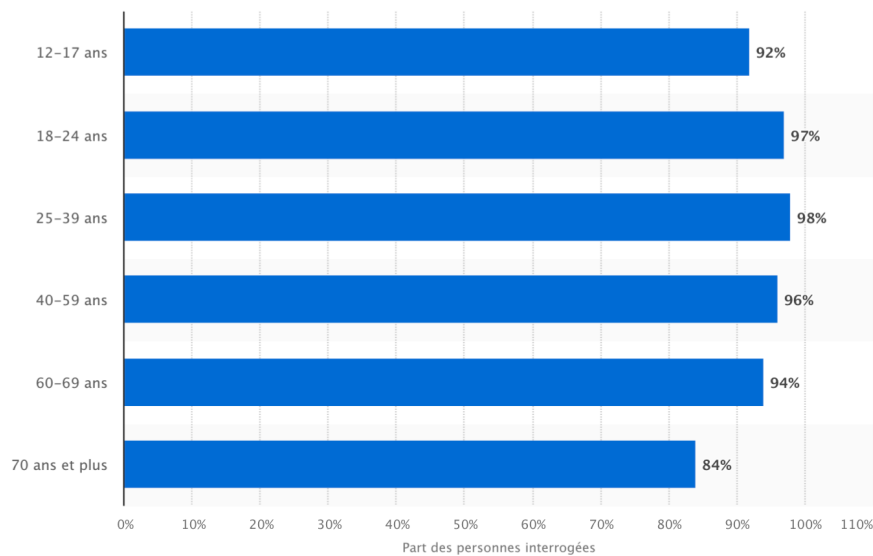
B.1.4. Creation of data

Figure B.9.: Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2020, with forecasts from 2021 to 2025 [196].



B.1.5. Smartphone equipment

Figure B.10.: Rate of smartphone ownership in France in 2020, by age group. [35]



B.1.6. Time spend on smartphone

Figure B.11.: Average time spent on a smartphone in February 2021 in the U.S. for people aged 18 and older. [30]

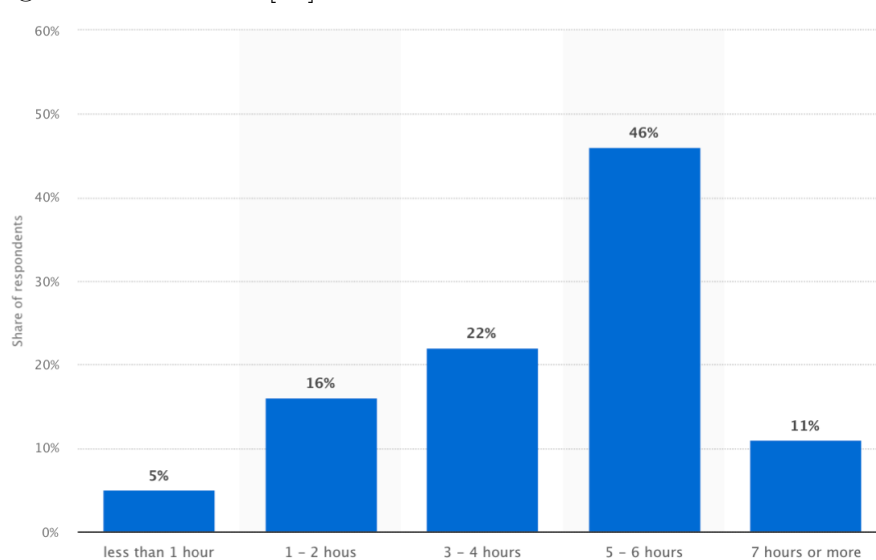
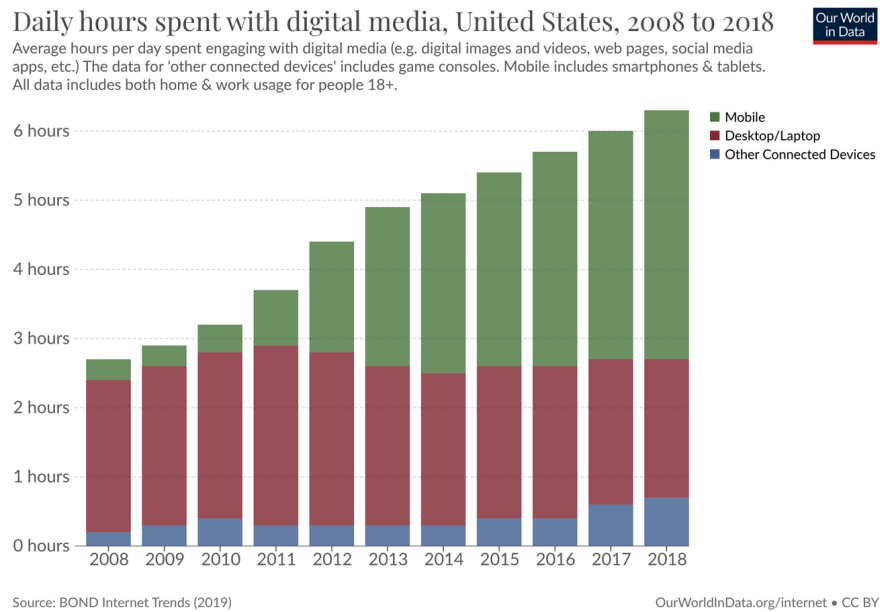


Figure B.12.: Daily hours spent with digital media, United States, 2008 to 2018. [282]



B.1.7. Most visited websites

Figure B.13.: Most visited websites in 2022 (in billion). [112]

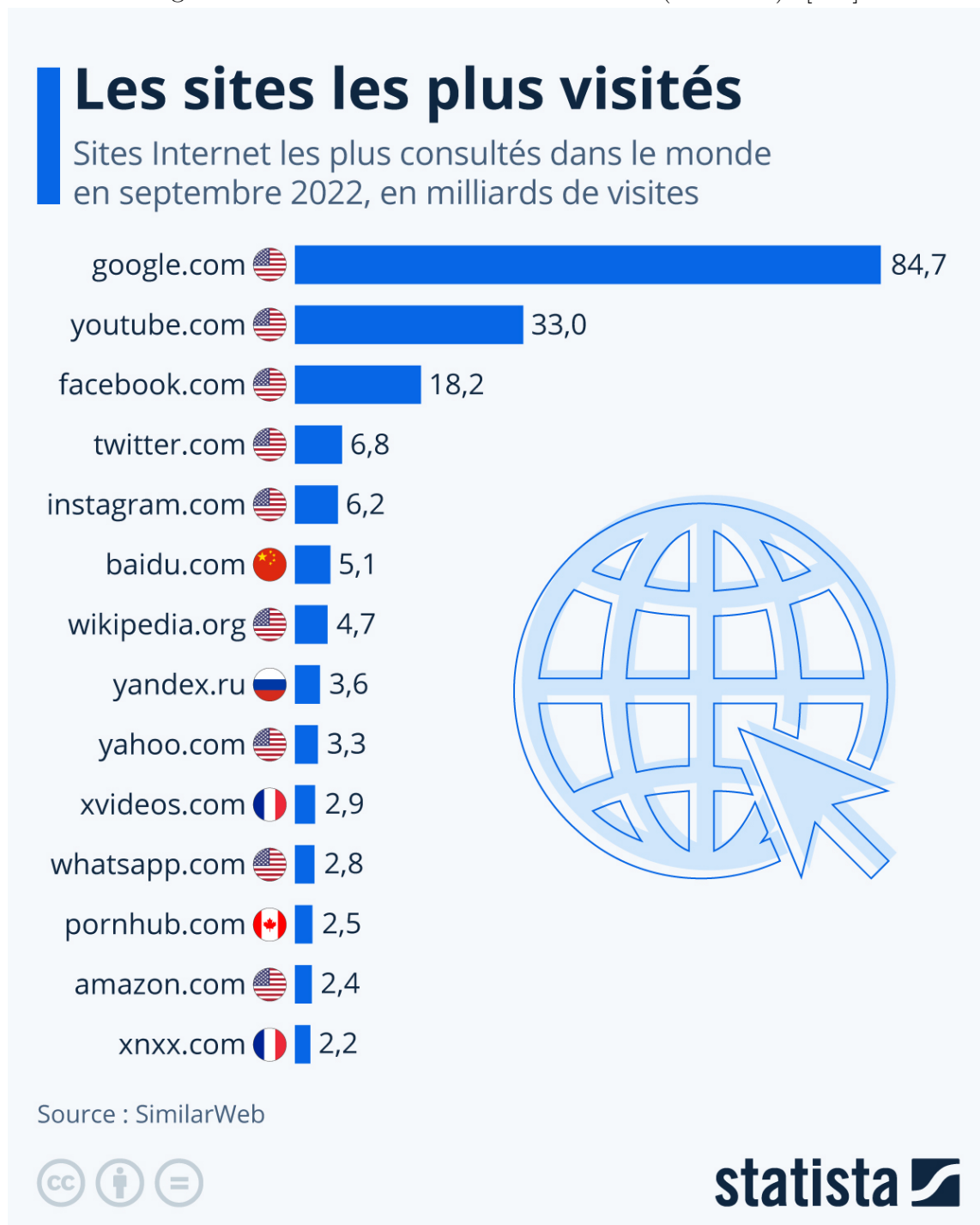
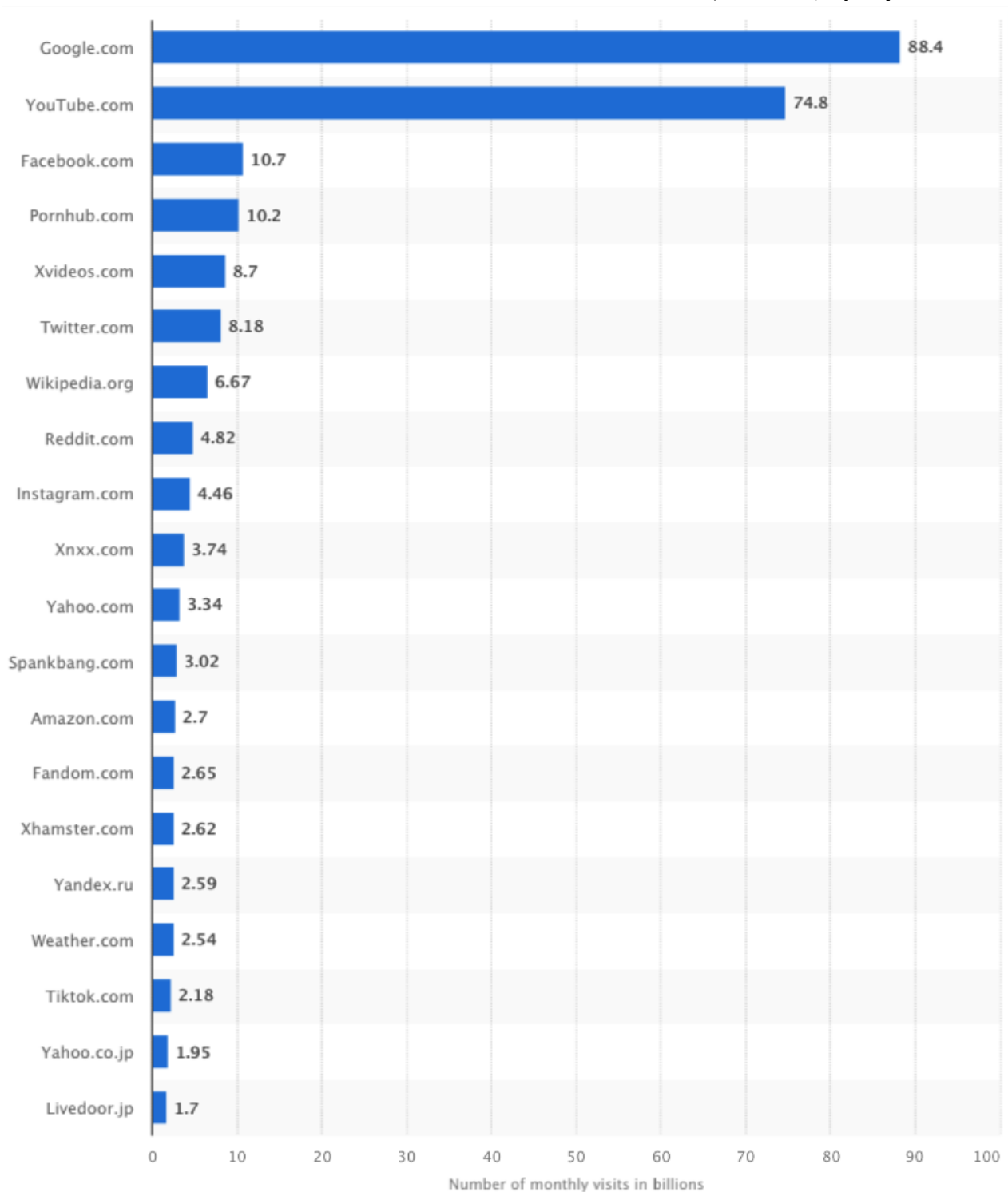


Figure B.14.: Most Visited Websites by Traffic in the world for all categories, January 2023.
[11]

Most Visited Websites by Traffic in the world for all categories, January 2023

Position	Website	Change	Visits	Pages / Visit	Bounce rate
1	 google.com	–	102,240,533,524	3.51	41.04%
2	 youtube.com	–	86,866,891,068	5.56	33.24%
3	 facebook.com	–	12,702,191,083	2.72	52.39%
4	 twitter.com	–	8,902,150,430	1.72	72.85%
5	 wikipedia.org	–	7,722,191,738	2.16	60.22%
6	 instagram.com	↑ 1	5,602,372,440	2.04	66.18%
7	 reddit.com	↓ 1	5,379,819,327	2.83	54.84%
8	 asurascans.com	↑ 12	4,331,036,419	5.29	33.25%
9	 lectortmo.com	↓ 1	4,136,835,014	5.19	51.59%
10	 fandom.com	↑ 3	3,265,085,704	3.27	50.72%
11	 yahoo.com	↓ 1	3,207,091,891	3.13	46.05%
12	 weather.com	↓ 3	3,116,264,834	1.50	72.81%
13	 amazon.com	↓ 2	2,891,171,662	5.47	43.80%

Figure B.15.: Most visited website in november 2022 (in billion). [112]



B.1.8. Users by social networks

Figure B.16.: Number of people using social media platforms, 2004 to 2018. [154]

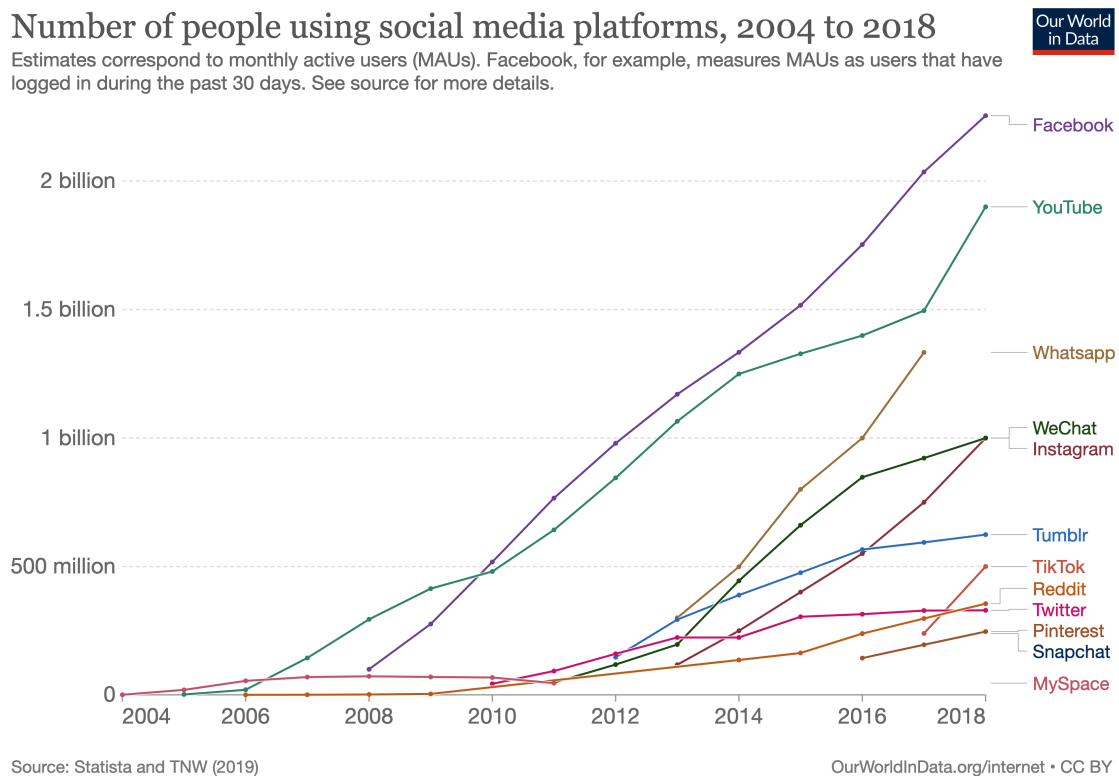
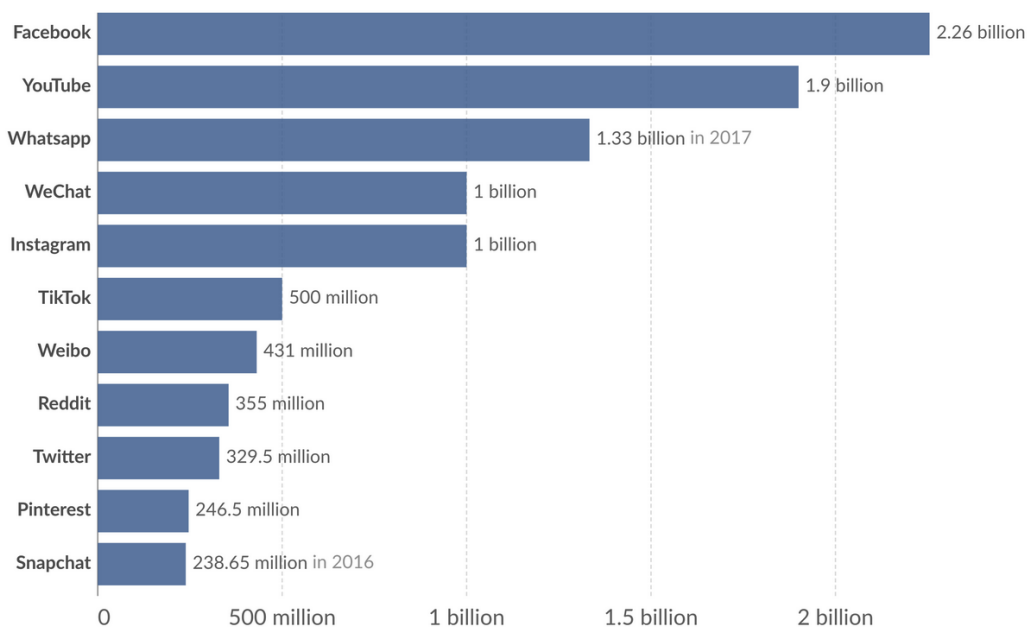


Figure B.17.: Number of people using social media platforms, 2018. [282]

Number of people using social media platforms, 2018

Estimates correspond to monthly active users (MAUs). Facebook, for example, measures MAUs as users that have logged in during the past 30 days. See source for more details.



Source: Statista and TNW (2019)

OurWorldInData.org/internet • CC BY

B.2. How to become a ghost on the web?

The title of this section sells dreams. It is very difficult to disappear from the Internet and to be incognito while we surf on the web. The goal with this project is not to become a ghost but that all the accounts we will create seem real and be considered as different people.

So a very important and crucial point in this project is firstly that I am not found to be the source of these created accounts. Then, that all these accounts are considered as different persons. The task is not as simple as it seems. I will explain why in this section and explain how we will proceed to be incognito.

B.2.1. The usefulness of a VPN

A VPN or Virtual Private Network is a service that establishes a secure, encrypted connection between your computer and the Internet. By doing so, you benefit from a private tunnel for your data and communications when surfing on public networks. Rather than sending your Internet traffic (online searches, uploads, downloads...) directly to your ISP, the VPN routes your traffic through its VPN server first. This way, when your data is finally transmitted over the Internet, it appears to be coming from the VPN server and not from your personal device [12] [7]. So it is a VPN server that makes the requests for us and sends us back the answers of the servers contacted for our requests.

Thus, with a VPN, we appear to be located in the country where the VPN server makes the request for us, because it makes our requests for us. It's interesting if we want to make believe the server that we want to contact (for example : Facebook) that we are in the USA while we are in Belgium. But it's not enough to be incognito as I will explain you in the Appendix B.2.2.

B.2.1.1. Won't I be suspicious by using a VPN for Facebook ?

Not at all ! VPN services are becoming more and more popular and have more and more users. By using a VPN and especially a popular VPN Service, I will blend in and not be suspicious for Facebook as many of their users use a VPN.

To give you some figure, NordVPN had more than 10 million subscribers in 2020 [18]. Another article say that there are 5 billion users of the Internet and among these users there is 1,2 billion users that use a VPN in 2022 but beware, 70% of these VPN users use free VPNs which are not always 100% reliable [260]. Lastly, another article say that 28% of users of Internet use a VPN in the world. India is the biggest consumer with 43.2% of Internet users using a VPN. In France, Belgium and Italy, we are around 20% [202].

Using a VPN is therefore a good idea to locate ourselves in another country that the service we use allows us. As there are many users of these services, we are not suspected of using one.

B.2.1.2. VPN that we will use

There are many good VPNs services. Here is a list in decreasing order in terms of excellence in 2022 although they are all more or less equivalent, especially for the top 3 :

- ExpressVPN
- CyberGhost VPN
- NordVPN
- Surfshark VPN
- Private Internet Access VPN
- Avast VPN
- ProtonVPN
- IPVanish
- Hide My Ass
- PureVPN

We have decided to use NordVPN to start this project because the price is not too high and it offers good functionalities. It allows us to connect in 59 countries, the connection is stable and fast. If we need, we can turn to other VPN services such as ExpressVPN or CyberGhost VPN to expand our range of countries where we can connect. For the moment, NordVPN is sufficient.

B.2.2. Browser fingerprint

Have you ever heard about browser fingerprint ? If not, it's time to find out what it is because websites have found a (very efficient) way to track us without using cookies. Moreover, no law forbids this system (even in Europe, which is one of the continents that does the most to protect its citizens). I discovered this by chance while doing research and I will explain the principle and a possible solution to protect yourself even if it is almost impossible. So first, I will explain you what it is and then, our solution to try to bypass it.

B.2.2.1. What is a browser fingerprint ?

This concept is based on 2 things, the APIs of your browser, javascript, HTML header (so information accessible via your browser) and personal information about your computer.

In 2010 (so it's a fairly recent concept) [95], Peter Eckersley of the Electronic Frontier Foundation revealed the possibility of exploiting browser fingerprinting to track Internet users. To illustrate this phenomenon, he launched the site <https://panopticlick.eff.org/> on which he installed a very simple script that retrieves 8 attributes of the browser and the environment of the site visitor. He showed that 94% of the 500,000 fingerprints collected at the time of the study were unique, which could be used to track Internet users. It also demonstrated that it was possible to track the same fingerprint over time by identifying simple changes. There are different fingerprints which, if combined, make us, with a very high probability, unique : canvas fingerprinting, WebGL fingerprinting and rendering fingerprinting, Device fingerprinting, ... [22].

B.2.2.2. How it works ?

As you can see, the fingerprint is based on the information and specifications of your computer and browser. One may ask how can they identify me only with that? The answer is simple, the more you use your computer, the more your computer will contain elements that uniquely identify you. This can be the type of fonts your computer has, the plug-ins you installed, the extensions of your browser, the version of your browser and OS, the type of your OS, the language of your computer and browser, your IP address, the size of your screen, the color palette of the screen, By collecting all these attributes, and making a hash, you are 90-99% uniquely identifiable on the web.

The list of fonts and plugins is often the most discriminating attribute in a fingerprint [95]. Indeed, when you receive a brand new computer, you have a font base. But depending on your use, you will install and uninstall software. When you install a software, it can add new fonts to your computer and conversely, when you uninstall software, it can remove them. Therefore, the more you use your computer, the more you contribute to having a unique combination of fonts installed on your computer. This information combined with several other attributes allows us to identify you easily on the web.

To give you a concrete example, I asked to my friends the number of fonts that they have on their computer (with the OS Windows). Here are the numbers they gave me : 282, 184, 195, 191, 192, 198, 186, 235, 189. It's good to know that personally I reinstalled windows recently and I have 81 fonts. So the more you use your computer and install/uninstall software, the more your computer differs from others ! On Mac, I noticed the same thing on the computer of my friends. Also, looking at my friends' fonts, if one person A had 250 fonts and another person B had 150 fonts, it doesn't mean that person A had all the fonts of person B. So you quickly realize that there is a huge amount of possible font combinations. This proves us that each person is different (based on their computer) without necessarily being aware and it just due to their use of their computer. By combining various attributes, it is easy to see that it becomes easy to identify us in a unique way on the web. If you want to see if you have an unique fingerprint, you can see it with the web site <https://coveryourtracks.eff.org/> that compare your fingerprint with all the fingerprint they collected since 45 days or <https://amiunique.org/> that compare your fingerprint with all the fingerprint they collected in the last 2 years. Those two website use different techniques to build your fingerprint. I tried it with my Mac on both websites and I'm effectively unique compared to the fingerprint they have in their database (212 064 fingerprints for the website <https://coveryourtracks.eff.org/> and 861 361 fingerprints for the website <https://amiunique.org/>). In addition, they both give a lot of feedback on our fingerprint and the website <https://amiunique.org/> indicates what proportion of users share the same items as me (see figure B.18). There are also more details on the page that I don't show in this paper. We can see that in each item, I'm not unique because other users have the same configuration as me. But by combining those items, I'm the only one with this combination.

Figure B.18.: Proportion of users sharing the same items as me.



To give you another example, in the article [95], they explain they have coded a fingerprint script on the site <https://amiunique.org/> where they get 13 attributes of the browser and its environment. Here are the principle sources of information:

- The header of the HTTP that send the client
- Script (like javascript) that used the API of the browser to get information about the system of the user.
- Plugin used by the browser.

Here is an example of information (figure B.19) they get when you launch they script on <https://amiunique.org/>.

Figure B.19.: Attributes with how they retrieve it and an example of values used by the website <https://amiunique.org/> to create a fingerprint

Attribut	Source	Valeur
User agent	En-t��te HTTP	Mozilla/5.0 (X11; Linux x86 64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/ 41.0.2272.118 Safari/537.36
Accept	En-t��te HTTP	text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,*/*;q=0.8
Content encoding	En-t��te HTTP	gzip, deflate, sdch
Content language	En-t��te HTTP	en-us,en;q=0.5
Liste des plugins	JavaScript	Plugin 1: Chrome PDF Viewer. Plugin 2: Chrome Remote Desktop Viewer...
Fuseau horaire	JavaScript	-60 (UTC+1)
Do Not Track	En-t��te HTTP/JavaScript	Oui
R��solution d��cran	JavaScript	1920x1200x24
Liste des polices	Flash	Abyssinica SIL,Aharoni CLM,AR PL UM- ing CN,AR PL UMin HK,AR PL UMin TW...
Plateforme	Flash	Linux 3.19.1-201.fc21.x86 64
WebGL Vendor	JavaScript	NVIDIA Corporation
WebGL Renderer	JavaScript	GeForce GTX 650 Ti/PCIe/SSE2

Another site <https://coveryourtracks.eff.org/> also allows you to get your fingerprint

according to the attributes they get and it's very sophisticated and clever. If you want to have more information, get a look at this link <https://coveryourtracks.eff.org/about> and at the Methodology section. When it comes to tracking us, sites do not lack imagination and ingenuity !

B.2.2.3. Is it really used by websites ?

Of course ! It is an incredible source of information to identify us on the web. In 2013, researchers from the KULeuven (in Belgium) have analyzed the 10 000 most popular websites to determine the propagation of the fingerprint on the Internet. They explain that websites use it more than we think. They also identified large commercial companies involved in fingerprinting, a complete disregard towards the DNT header, and the use of anti-debugging techniques, most commonly associated with JavaScript malware. Moreover, they showed that dedicated fingerprints can bypass existing privacy-protecting technologies [54].

Moreover, companies like Google, Yahoo, twitter or Amazon do not indicate in their privacy policy the use of fingerprinting, but they do state that they collect data about the device used to improve the services offered to users [95].

So, it's certain that these giant companies use the fingerprint, it's all in their interest and it allows them, in Europe, to continue to track the user easily because Europe puts a lot of restrictions on the use of cookies and the tracking of its citizens by companies with the ePrivacy Directive [40] [248].

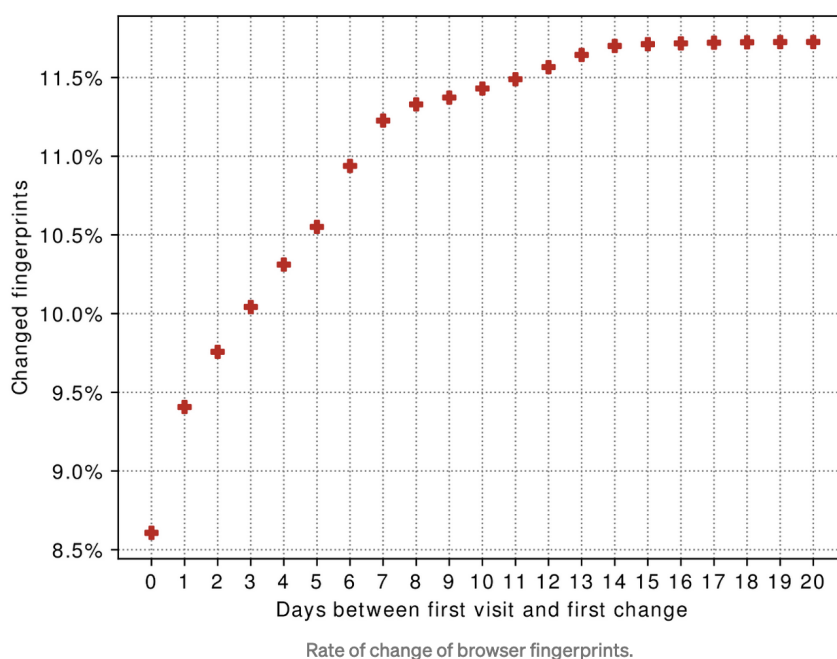
B.2.2.4. Yes but my fingerprint can change over time according to my use of my computer. So I'm not longer traceable ?

That's true that you fingerprint can change according to the use you have with your computer ! If you install a new plug-in or software, your fingerprint hash may change (depending on the attribute the websites get). But we must not claim victory due to two factors.

First, we don't everyday install, uninstall software or plugin. We may install a lot of new things when we buy a new computer but when all our software are installed, we modify very little our computer (at least the majority of people). This is less the case for people working in IT (but this is a minority of users). I can give you the example of my parents. They have the

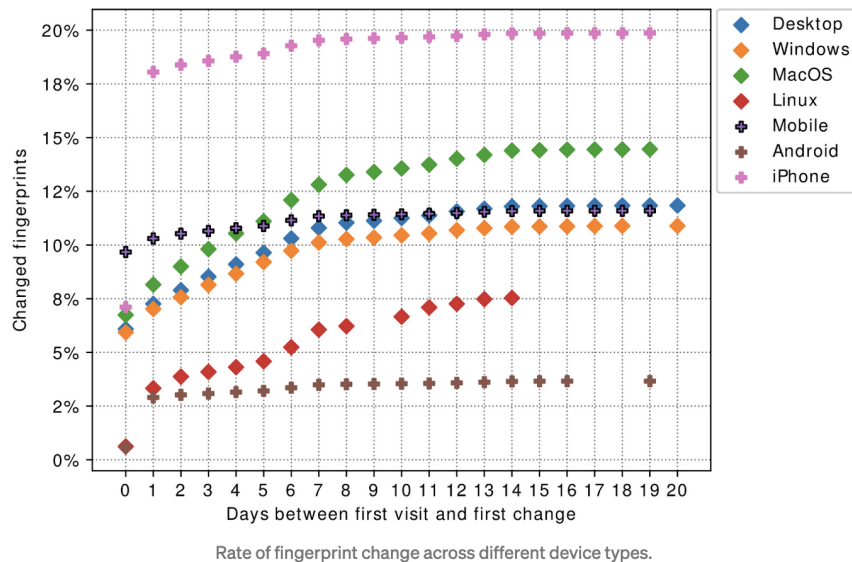
same computer since 2010 and they never download something. They just use their computer to look at their mails, books vacations, makes research and to make their accounting. To give you a concrete example, there is a study that studied how fast our fingerprint changes (and other things) on 566 704 browser fingerprint (65% of the devices was smartphone) [289]. They give us an interesting graph (figure B.20).

Figure B.20.: Number of days between the devices first visit and the first change of its browser fingerprint. [289]



They also give a graph for each type of device separately (figure B.21).

Figure B.21.: Number of days between the devices first visit and the first change of its browser fingerprint depending on the type of the device. [289]



We can see on figure B.20 that after 20 days, only 11,5 - 12% of the devices have another fingerprint. And the curve seems to no more increase. On the figure B.21, we see that the brand Apple seems to have more protection for that because the curve of the Iphone and MacOS are the highest ones.

Second, if your fingerprint changes, they will very quickly re-recognize us and as they track us in other ways with other tools, they will after some time know that our fingerprint has changed. It also demonstrated that it was possible to track the same fingerprint over time by identifying simple changes as said previously [22]. In any case, even if our fingerprint changes, it will remain the same for a sufficient time to track us until it changes again.

B.2.2.5. Why browsers allow web site to get those information ?

It is not the purpose of browsers to give this information to the websites. Over time, browsers have improved by allowing users to customize their browser (plug-in, color, ...) and the developer to offer dynamic web content, attractive and adapted to the environments of Internet users (for example, the site is automatically displayed in the language chosen by the user).

The appearance of browser fingerprinting as a tracking technique is a direct side effect of

When a web site wants to get the different types of fonts we have, it is on the one hand legitimate because it allows him to know with which font the text will be displayed on our screen but on the other hand, it has a harmful effect because the site can use it to identify us. It is therefore not possible to know for what purpose a web site requests this information.

When company want to generate new data, it's not difficult for them. Here is an example.

221

B.4. Experiments conducted by Mr. B.J. Fogg

Here is a summary of the experiments. Don't hesitate to consult this source [161] which is more detailed and contains all the experiments.

The experiments were carried out with samples of 50 to 75 people. The main results are following:

- A first experiment aimed to give a personality to the computer on the way the popup messages were written in order to create dominant or submissive computers. The 48 participants (selected according to their personality (the most dominant and the most submissive personalities)) had to solve the Desert Survival Problem and received instructions and recommendations from the computer. The information given by the computers was all the same but differed in the way it was said. At the end of the experiment and after surveying the participants, they came to the conclusion that people prefer computers that match their personality. Therefore, if a computer resembles us, it is more likely to persuade us.
- Second experiment (on the principle of similarity) and even more interesting is the following: 56 participants had, again, to solve the Desert Survival Problem. Here, all the computers were similar and gave the same instructions and recommendations without any difference in the way they were written. The difference here was that he told each participant, before they started solving the problem, whether the computer was his teammate or just a simple computer (without a label). The conclusion was that: people who worked with a computer labeled as their teammate reported that the computer was more similar to them, that it was smarter, and that it offered better information. These participants were also more likely to choose the problem solutions recommended by the computers.
- The last experiment carried out is on the principle of reciprocity. The 76 participants had to solve a modified version of the Desert Survival Problem where they needed to perform searches on a computer. Half of the participants had a computer that gave very useful answers to their research and that stipulated to search in several databases to find the best answer (in order to give the feeling that the computer is doing us a favor and is trying to find an answer) while the other half had a computer that gave information of lesser quality and not always useful. At the end of the experiment, the participants

had to help the computer (if they wanted to) to create a color palette that matched the human perception. The computer gave 3 colors and they had to classify them from the lightest to the darkest. The participants could do as many iterations as they wanted. They noticed that those participants who worked with the same helpful computer on both tasks performed almost twice as much work for their computers on the second task as did the other participants.

B.5. Cambridge Analytica

B.5.1. The story behind the creation of Cambridge Analytica

Cambridge Analytica is a sub-branch of the British firm SCL Group (Strategic Communication Laboratories), which provides governments, political groups and corporations around the world with services ranging from military disinformation campaigns to social media branding and voter targeting [232] and was therefore created to shape public opinion and influencing behavior without being seen [146][197]. One of the founders (Mr. Nigel Oakes) had the objective, before the creation of this company, to use psychology to manipulate people and make them behave as the client wanted [146]. Mr. Oakes explained in 1992 to the newspaper *Marketing*, that they used the same technique as Aristotle and Hitler. That is, We appeal to people on an emotional level to get them to agree on a functional level [306].

The company was awarded a contract in 1999 to enhance the reputation of Indonesian President Abdurrahman Wahid [197][306]. They also had government contracts, including a Foreign Ministry contract to counter violent jihadist propaganda in Pakistan. They have had more and more contracts including one in 2015 to help Nato states counter Russian propaganda in eastern Europe. In parallel to their military work, they turned to politics to help politicians win by manipulating people [197]. See figure B.22 to have an overview of the structure of the firm SCL group and all its sub-branch.

Concerning Cambridge Analytica, all started in 2013 when Mr. Alexander Nix (who was a director of the SCL Group) met Mr. Wylie who was interested in using inherent psychological traits to affect voter behavior and had found a team of psychologists to help him [318][124].

Subsequently, Mr. Nix had the good fortune to meet Mr. Steven Bannon⁴, Mr. Robert Mercer who who believed a sophisticated data company could make him a kingmaker in Republican politics and his daughter Rebekah (who would become a board member) and who shared his conservative views [318]. After a meeting, Mr. Bannon and Mr. Mercer accepted to invest in the creation in Cambridge Analytica. Mr Mercer invested \$10 million [191].

In 2013 the testing began and Mr. Mercer injected \$1.5 million to survey voters and test psychographic messages in the Virginia gubernatorial race where Republican Attorney General Ken Cuccinelli ran against Democratic fundraiser Terry McAuliffe. Although Monsion Cuccinelli lost, Mr. Mercer wanted to continue the search and wanted results [318].

In early 2014, investor Toby Neugebauer and other wealthy conservatives wanted to invest tens of million of dollars to help Texas Senator Ted Cruz win his presidential campaign [318]. In the end, at least 750,000 dollars were paid by Mr. Cruz's team and more than \$11 million from Mr. Mercier (who supported Mr. Cruz [290]) for the pro-Cruz super PACs [181][268]. In the meantime, Cambridge Analytica also worked, that same year (2014) with Mr. Ben Carson by receiving 220,000 dollars[181]. Finally, although Cambridge Analytica's techniques were not yet perfected, Mr. Cruz won the Iowa and Texas state elections. The problem was that Cambridge Analytica was running out of data and that is when Mr. Wylie contacted Mr. Kogan. At that time, Mr. Mercer agreed to invest \$15 million in a joint venture with SCL's elections division. In the summer of 2014, Cambridge Analytica embarked on the mid-term elections and mobilized contractors and SCL employees across the country [318].

As we can see, Cambridge Analytica was composed exclusively of investors and politicians of extreme right and thirst for power. The goal was clear: for the politicians, it was to manipulate people to vote for them and for the investors the goal was to increase their power even more. The pepper of money was not enough for them, they wanted to have the power over the population by controlling and manipulating them through the data they generate, mathematical models and advertising.

⁴He is a conservative activist close to the extreme right and was the executive chairman of the conservative newspaper Breitbart News since 2012 [208][294]. He then became Trump's presidential campaign manager and White House adviser. Regarding Cambridge Analytica, he became its vice president.

B.5.2. Brexit

Cambridge Analytica is a sub-branch of the firm SCL group (see more explanation in Appendix B.5.1) which is based in England [350][290][223] and had contracts with political parties. The firm SCL group tried to influence the vote in favor of Brexit [290]. Indeed, Mr. Nix (he was a director of the SCL Group) wrote himself an article the 10 February 2016 in the Press Campaign where he described how his company (Cambridge Analytica) had helped supercharge Leave.EU's social media campaign by ensuring the right messages are getting to the right voters online. [65][197].

Although he clearly explains that they tried to manipulate the population, there are still doubts about the effectiveness of their methods. Frederike Kaltheuner of Privacy International explained that when testing their technology on its own data, the results were not very conclusive [197].

To this, we can also add the fact that at the launch of Leave.EU, Brittany Kaiser, an employee of Cambridge Analytica, appeared on stage for Nigel Farage's pro-Brexit campaign group and claimed to have started interviewing the British. Moreover, in March 2017, Mr. Arron Banks, a Bristol-based insurance entrepreneur who helped fund the group, wrote on Twitter: *We didn't hide that we were working with Cambridge*, and then denied (after losing the official designation to Vote Leave) that the SCL Group had played a major role in the Leave.EU campaign after the election [306].

It is Vote Leave (which included Boris Johnson) that has obtained the official designation. During their campaign, they paid a little more than 3.5 million pounds (more than 50% of the budget of their campaign) to a company founded in 2013 and named Aggregate I.Q (which specializes in targeted Facebook advertising and profiling in Canada) [197]. If we dig a little deeper into this story, we realize that Aggregate IQ and the SCL Group are closely linked. Indeed, first, there was an intellectual property license that linked Aggregate I.Q and SCL Groupe that attributed the intellectual property of Aggregate I.Q. to Cambridge Analytica. This made Aggregate I.Q. an internal department of Cambridge Analytica. Second, until February 2017, on SCL Groupe's website, their phone number for their office in Canada was that of Aggregate IQ [197][306].

We realize that by making the right connections, everything is directly linked to the SCL Group. Moreover, the two groups Leave.EU and Vote Leave seemed, at the base, distinct but

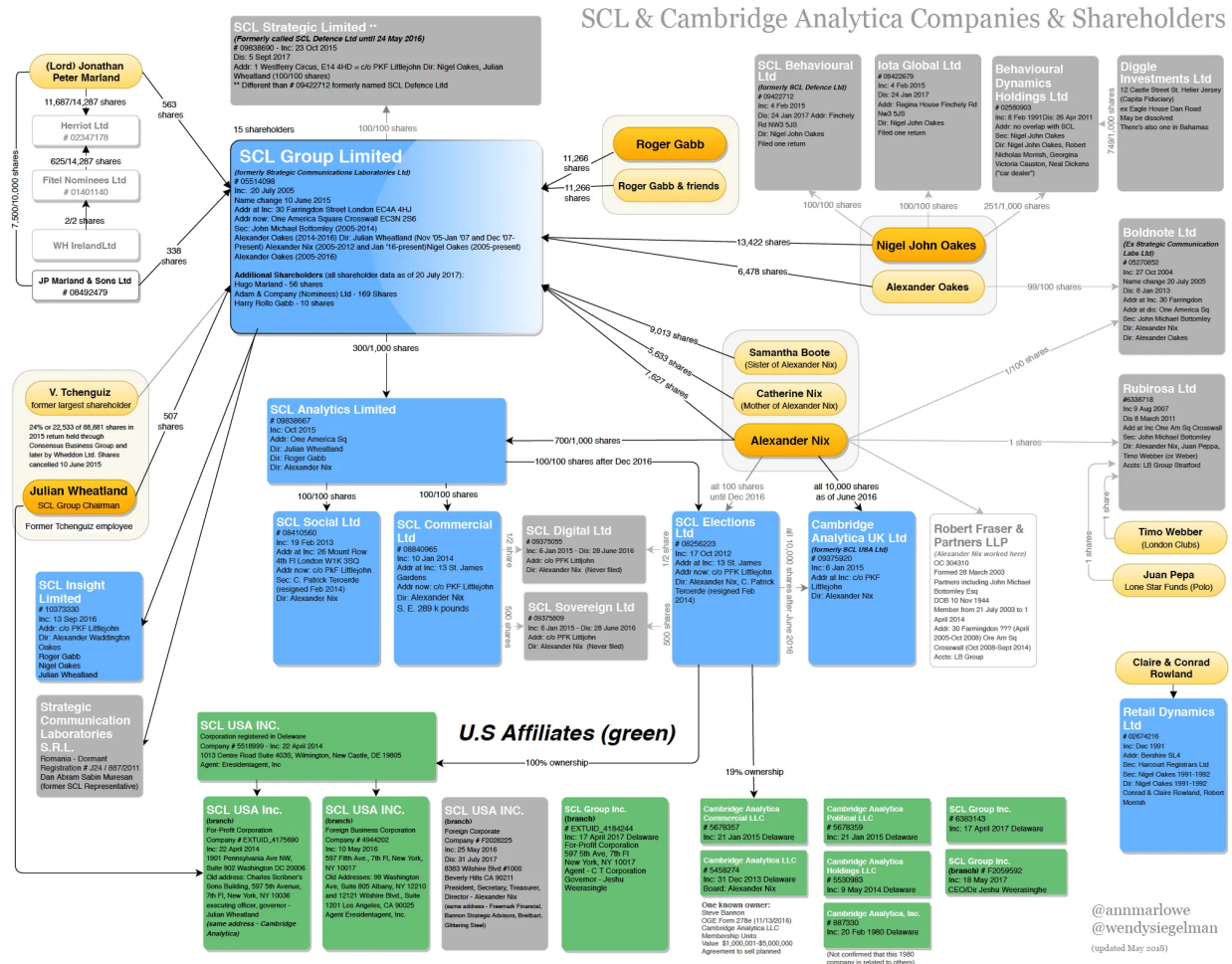
finally called upon the services of a single group: the SCL Group [101]. In our opinion, and based on the various documents we have consulted, we affirm that the SCL Group (parent of Cambridge Analytica) has influenced the population of England to adopt the Brexit. To have a visual diagram, a graph is given in the Appendix B.5.6 with all the connections between the different companies and political parties.

B.5.3. Russia

To finish the story of Cambridge Analytica and SCL Group, it is important to know that they also had a relationship with Lukoil (Russia's second biggest oil producer). Although the SCL Group and Lukoil deny that their conversation was political and say that Lukoil was never their client, it is still interesting to ask why Lukoil contacted them? [267]. The reason was that they wanted to understand how their services are going to apply to the petroleum business because they understand behavioural microtargeting in the context of elections but that they were failing to make the connection between voters and their consumers [98]. The SCL Group's response was to be shared with the CEO of Lukoil and Putin's associate: Vagit Alekperov. There is no evidence that Cambridge Analytica worked for Lukoil but it is clear that some documents show that in 2014, Lukoil was fully briefed on Facebook, microtargeting, data and election disruption [98].

B.5.4. Structure of SCL & Cambridge Analytica Companies & Shareholders

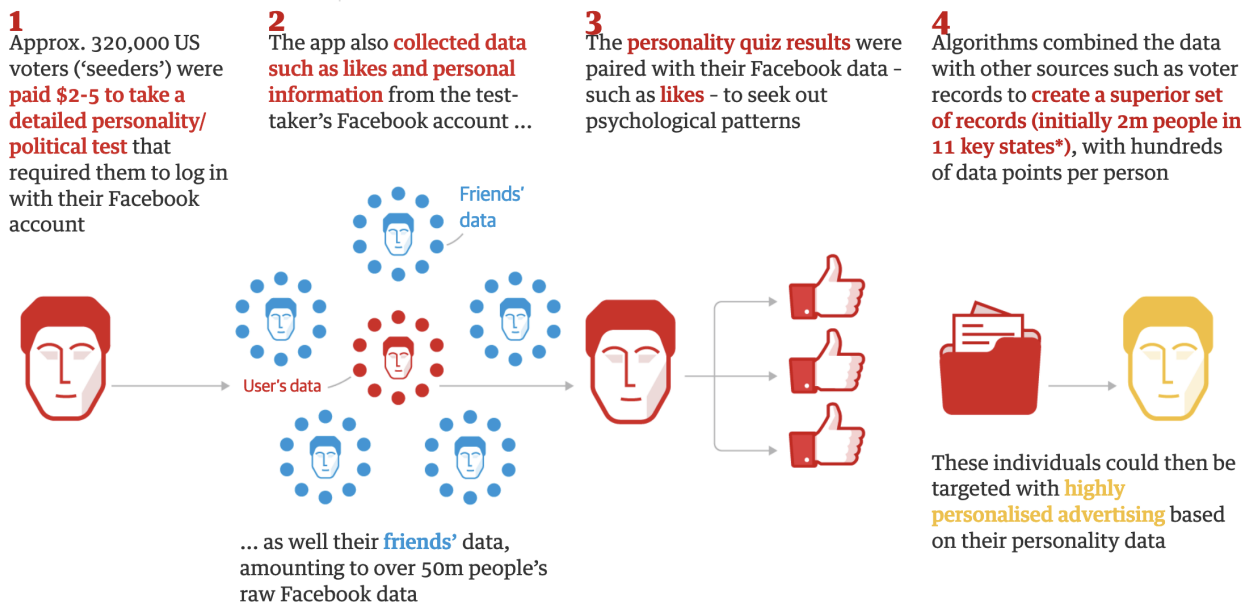
Figure B.22.: Structure of SCL & Cambridge Analytica Companies & Shareholders. [350]



B.5.5. Cambridge Analytica - graphs

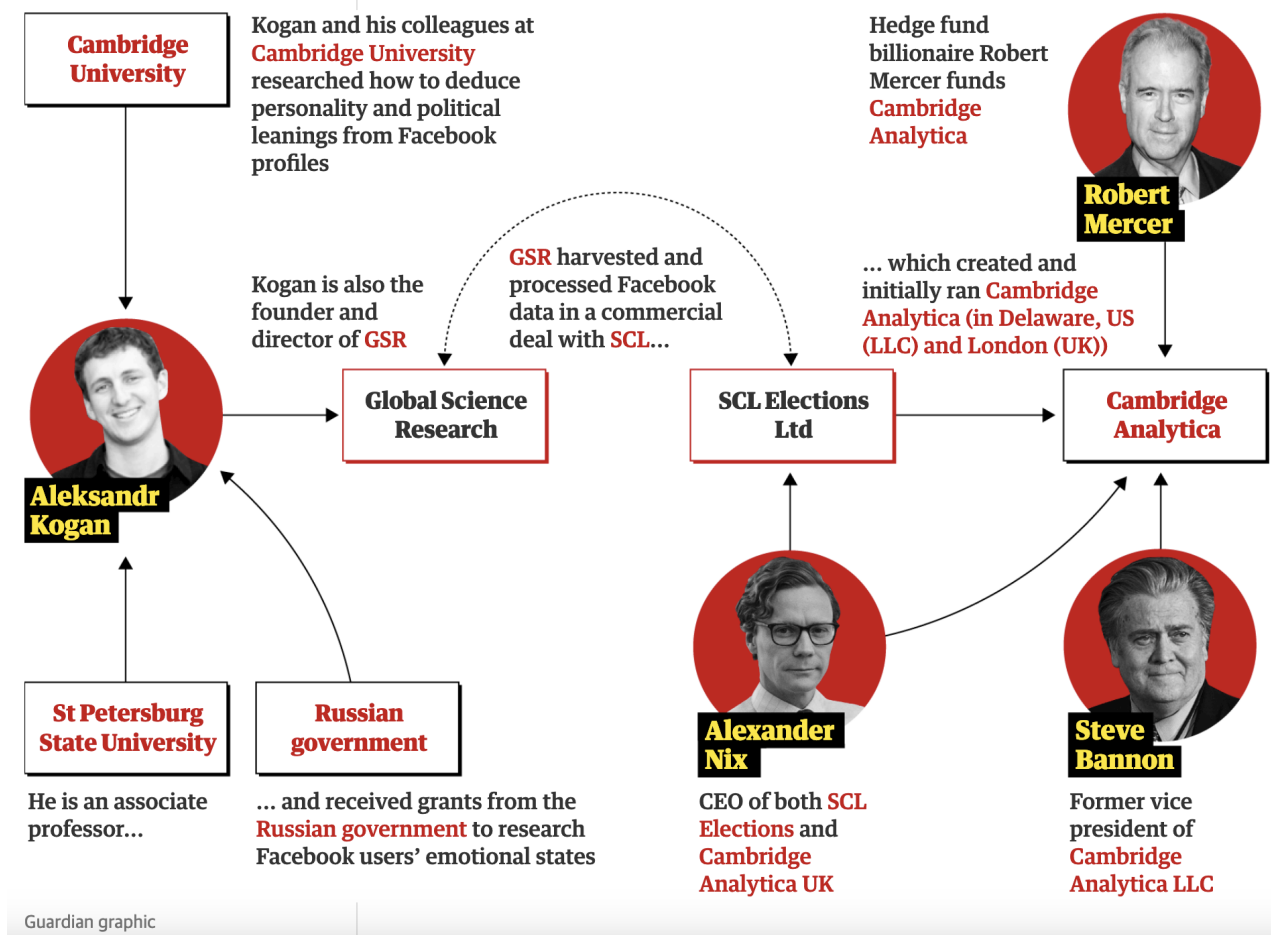
Figure B.23.: Cambridge Analytica: explanation of how 50m Facebook records were hijacked. [99]

Cambridge Analytica: how 50m Facebook records were hijacked



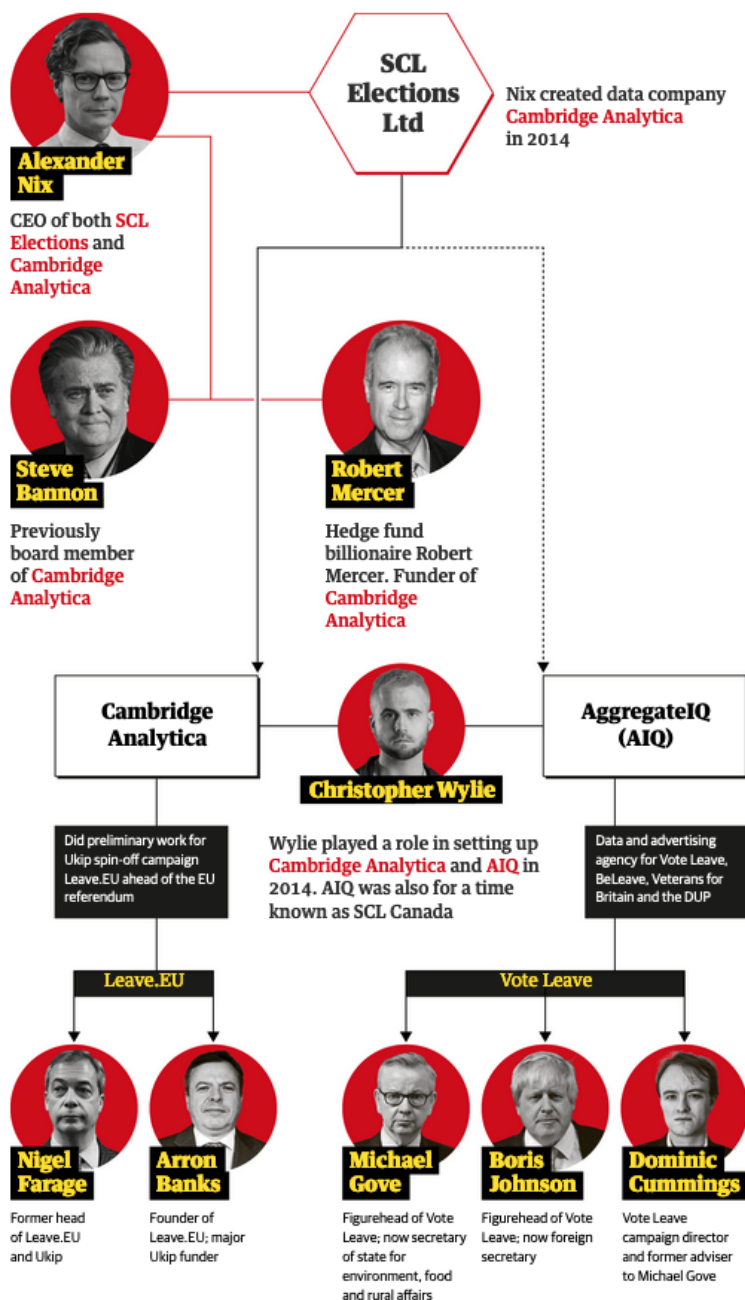
Guardian graphic. *Arkansas, Colorado, Florida, Iowa, Louisiana, Nevada, New Hampshire, North Carolina, Oregon, South Carolina, West Virginia

Figure B.24.: Cambridge Analytica: how the key players are linked. [100]

Cambridge Analytica: how the key players are linked

B.5.6. Cambridge Analytica - Brexit

Figure B.25.: Worldwide web: the data companies behind Brexit. [101]

Worldwide web: the data companies behind Brexit

B.5.7. Cambridge Analytica - Example of advertisements

B.5.7.1. Example 1

Figure B.26.: A campaign advert tested by Cambridge Analytica. [60]



B.5.7.2. Example 2

Figure B.27.: A campaign advert tested by Cambridge Analytica. [60]



B.5.7.3. Example 3

Figure B.28.: Visual messages tested by GSR's online profiling test. Respondents were asked: How important should this message be to all Americans?. [98]

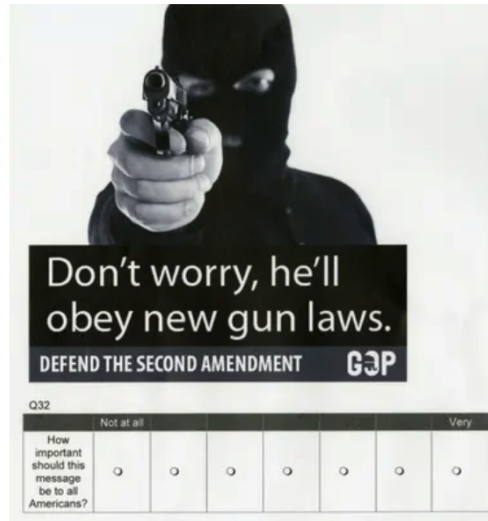
**B.5.7.4. Example 4**

Figure B.29.: Visual messages tested by GSR's online profiling test. Respondents were asked: How important should this message be to all Americans?. [98]



B.5.7.5. Example 5

Figure B.30.: Visual messages tested by GSRs online profiling test. [98]



B.6. IRA

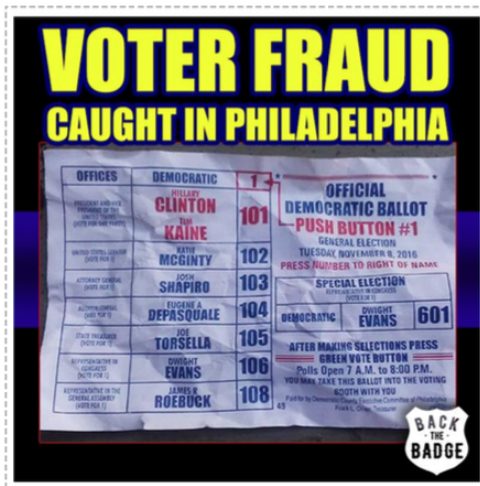
B.6.1. Example of advertisements and posts made by the IRA

B.6.1.1. Example 1

In the days leading up to the 2016 election, some Instagram accounts tried to cast doubt on the integrity of the election and accused the Democrats of trying to rig the vote in their favor.

The picture below show you a post of a conservative Instagram account which contains false accusation of voter fraud in Pennsylvania. It was also shared on Twitter and Facebook

Figure B.31.: Post that spread false accusation of voter fraud in Pennsylvania. [235]

Internet Research Agency Post**B.6.1.2. Example 2**

Several of the Instagram posts traced to Russia offered merchandise for sale. here, you can find a post shared by the @stay4police account who advertised T-shirts and hooded sweatshirts that read I Support American Law Enforcement. Other posts encouraged black Instagram users to support black-owned businesses with the hashtag #buyblack, and offered natural skin care products and other merchandise targeted at a black audience. [235]

Figure B.32.: Post shared by the account @stay4police which sell T-shirts and hooded sweat-shirts where we can read I Support American Law Enforcement. [235]

Internet Research Agency Post



Researchers explain that it was certainly not to make money that the IRA sold clothes but for two other reasons: (1) So that the buyers could encode their data and the IRA could collect their personal information. (2) To identify supporters of the cause so that they could target them with advertisements.[235]

B.6.1.3. Example 3

Figure B.33.: Posted on March 2, 2016 by @army_of_jesus_ received nearly 90,000 likes on Instagram, making it the most-liked Instagram content uploaded by the group before the 2016 election. [284][235]



B.6.1.4. Example 4

Figure B.34.: A post from an IRA-controlled Facebook page called Black Matters. [284]



B.6.1.5. Example 5

Figure B.35.: A post by the Kremlin’s Internet Research Agency mocking the idea of Russian interference in a post targeted to right-wing audiences. [284]



B.6.1.6. Example 6

Figure B.36.: A post from an IRA-controlled Facebook page called Secured Borders. [302]



222 Likes 19 Comments 106 Shares

B.6.1.7. Example 7

Figure B.37.: Posted by the @blackstagram_ account in June 2017. It received more than 250,000 likes and more than 6,000 comments. [235]



B.6.1.8. Example 8

Figure B.38.: The IRA pages offered coupons in exchange for sharing content.[132]



34301 @black4black
<http://www.chloeandchad.com/>
 "We are happy to present you our partner Chloe and Chad . These products are all made by hand; they are all natural and are made when ordered, so they are as fresh as possible. Using natural skin care is a great way to treat your skin right! Enjoy a good discount with our promo! Just DM us!
 (on FB: "Enjoy a good discount with our promo code after you share 3 posts from Black4Black or BM.")

B.6.1.9. Example 9

Figure B.39.: Cluster of Hillary Clinton memes, a cluster of Donald Trump memes, and other memes related to GOP politicians on the political page Being Patriotic. [132]



B.6.1.10. Other examples

Many more examples of post can be found at this source: [123]

B.6.2. The Top 20 IRA Facebook Pages

Here are the top 20 IRA Facebook pages sorted by number of like.

Figure B.40.: Top 20 IRA Facebook pages sorted by number of like. [185]

Table 5: The Top 20 IRA Facebook Pages, Sorted by Number of Likes

Facebook Pages Managed by the IRA (Top 20 Liked)	Number of Likes	Number of Shares	Number of Reactions	Number of Comments
Being Patriotic	6,431,507	4,429,880	399,542	393,179
Stop A.I. (Stop All Immigrants)	6,050,989	5,202,716	552,684	778,924
Heart of Texas	5,489,337	4,986,384	590,664	414,599
Blacktivist	4,642,946	4,843,595	1,411,605	509,882
United Muslims of America	2,479,294	1,268,022	265,716	175,976
Army of Jesus	2,359,018	651,106	262,113	387,765
Brown Power	2,098,769	1,300,998	373,643	128,795
LGBT United	1,974,368	930,199	396,257	87,500
South United	1,419,503	2,263,031	101,931	72,461
BM (Black Matters)	1,333,173	1,797,479	325,864	146,254
Secured Borders	1,220,079	713,905	121,553	117,824
Defend the 2nd	986,969	551,847	90,228	39,530
Williams&Kalvin	569,627	541,436	138,078	39,960
Woke Blacks	454,151	490,623	127,179	37,876
Back the Badge	410,364	155,524	63,765	26,274
Veterans Come First	330,662	307,021	45,057	33,302
Memopolis	135,704	78,996	21,061	13,002
Pan-African roots MOVE	124,938	152,931	44,929	15,655
Born Liberal	104,314	79,938	22,933	5,749
Black Matters	59,032	97,516	14,620	9,350
<i>Total</i>	<i>38,674,744</i>	<i>30,843,147</i>	<i>5,369,422</i>	<i>3,433,857</i>

Source: Authors' calculations based on data provided by the SSCL.

B.7. Europe and Facebook advertising

What about Europe? Well, first of all, Belgium is in the biggest spenders because we find 4 Belgian Facebook pages in the top 100 of politicals ads in Europe: NV-A, Vlaams-Belang, PVDA and Bart de Wever. Indeed, Bart de Wever is the highest advertised politician of Europe in 2021 [20]. At the time of the study (in 2021) Belgium occupied the 12th place of the countries spending the most on Facebook advertising per 1000 inhabitants. It was notably surpassed by countries that had had elections in 2020 (Netherlands, Iceland, Denmark, Finland) or that were going to have election (Sweden, Hungary) [20].

So in Belgium, it's the extremist parties that spend the most on advertising on Meta. But what about in other countries like Hungary (where there were legislative elections in 2022), Italy (where there were parliamentary elections in 2022), etc. ?

B.7.0.1. Hungary

On April 3, 2022, Mr. Viktor Orban won with 53.35% of the vote and will serve a fourth successive term as prime minister [87]. Mr. Orban is nationalist and ultra-conservative. He belongs to the far-right political party Fidesz and has a friendly relationship with Mr. Vladimir Putin [24][16].

Now let's look at Facebook. First, there is a political influencer network in Hungary called Megafon (which circumvents the rules on political campaign ads). This group was created in February 2021 and its founder Mr. Kovacs has close ties to the political party Fidesz and he is a member of the team behind the Budapest-based Center for Fundamental Rights that seem to play a key role in Fidesz communication [329]. Well, we can feel the same pattern as Cambridge Analytica in USA.... Their goal is to train personalities to pose as opinion leaders and spread messages to support Victor Orban and to denounce his challenger [345][10]. There is also the Aktualis Media network which also seems to be part of the same network because firstly, the general manager of Aktualis (Filep David) is a known member of the Megafon network and secondly, there has been spending since August 2021 that has been combined with messages very similar to that of Megafon which reinforces the idea that they are not independent entities [10].

When we look at the spending on Facebook, the top 10 political pages contain a lot of people who are members of the Megafon group. If we combine all these people, it represents around 650 000 euro spent on Facebook Ad during the month of December 2021 (3 months before the elections). That's half of the total expenses of all parties during this month [10].

B.7.0.2. Italy

On September 25 Georigia Meloni won the general election with her party Fratelli d'Italia (far right party) with more than 26% of the votes. She has maintained close ties with Mr. Viktor Orban (who congratulated her on his election in 2022) since 2018 [355] and congratulated Mr. Puttin on his re-election in 2018 [180]. Although she has potentially received support from both politicians, she defends and supports Ukraine against Russia [118][286].

What about advertising on Facebook? What emerges in general is that it is the far right parties that spend the most. The party Fratelli d'Italia (created in 2012) has had a significant

growth between the year 2018 and 2022 in terms of number of votes. In 2013 it had obtained 2% of the votes, in 2018: 4% of the votes and in 2022: 26% of the votes [171]. If we look at the money put into Facebook Ad, it is also the party that invests the most in this tool. Between the 22 August and the 20 September 2022, they had spent a little more than 140,000 euros in Facebook Ad [251].

B.8. Where do people get their news?

Here are different interesting documents realized by Pew Research Center which realized several statistics on the consumption of information on social networks in the U.S. in 2016 [206], in 2017 [205], in 2018 [228], in 2020 [70] and in 2021 [249].

In 2016, in the U.S., a majority of people (62%, i.e. 6 out of 10 Americans) got their news (not exclusively) on social networks (figure B.41), of which 44% get their news (not exclusively) from Facebook (figure B.43). Among these 62%, 1/3 get information very often via social networks (figure B.41). Moreover, 64% of social media news consumers get their news on only one site (figure B.44). We also notice that on Facebook, Instagram and Youtube, there is a majority of consumers who receive news by chance so it is the algorithm that proposed them without necessarily that it is their goal to receive this information (figure B.45). Compared to 2013, Facebook, Instagram, and Linkedin had the largest increase in the number of users getting information via their platforms (figure B.46). [206][68][230]

In 2017, the number of people who get information via social networks increased from 62% in 2016 to 67% in 2017 (figure B.47), of which 45% (not exclusively) on Facebook (figure B.49). The social networks that have had the most growth are Twitter, Youtube and Snapchat (figure B.48). However, it is interesting to note that people tend to diversify their source compared to 2016. Indeed, in 2016, 18% of Americans took information from 2 or more social media. In 2017, this has increased to 26% (figure B.50). [205]

In 2018, 68% of the population is informed by social networks. However, a majority of social network users find the information on social networks not accurate (figure B.51) and we are seeing a slowdown in the growth of information consumption on social media (figure B.52). [228]

In 2020 and 2021, we remain around 68% of people who inform themselves with social networks

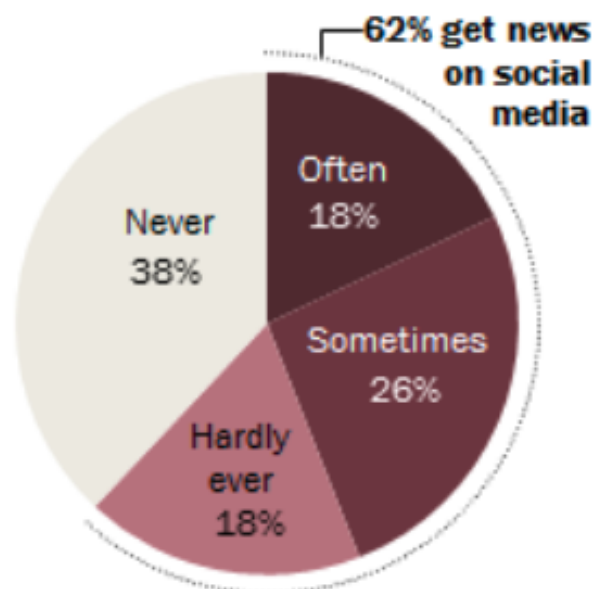
although we notice a generalized decrease (figure B.61). Maybe with the containment that was put in place to counter the covid-19, people have been using social networks so much that they are trying, now, to use them less. The scandal of Frances Haugen also contributed to questioning the population. Beware, the proportion of people informing themselves via social networks remains important! [70][249]

We also conducted our own survey at UCLouvain which you can find in the appendix D.1. In question 2/15, we see that almost 50% of students use social networks as their **main source**.

Figure B.41.: About 6-in-10 Americans get news from social media - 2016 [206]

About 6-in-10 Americans get news from social media

% of U.S. adults who get news on a social networking site ...



Source: Survey conducted Jan. 12-Feb. 8, 2016.

"News Use Across Social Media Platforms 2016"

PEW RESEARCH CENTER

Figure B.42.: Reddit, Facebook and Twitter users most likely to get news on each site - 2016 [206]

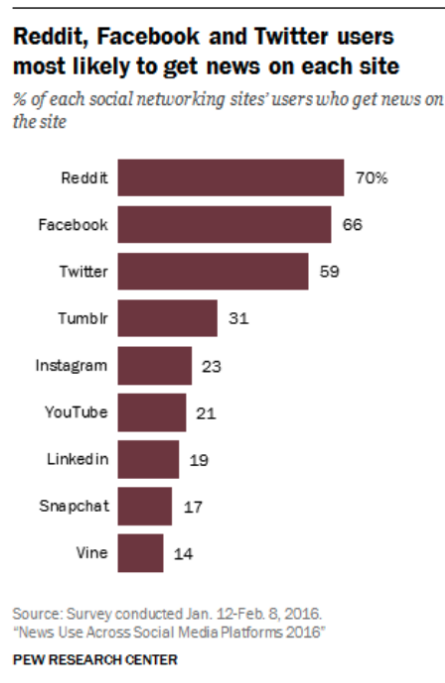


Figure B.43.: Social media news use: Facebook leads the pack - 2016 [206]

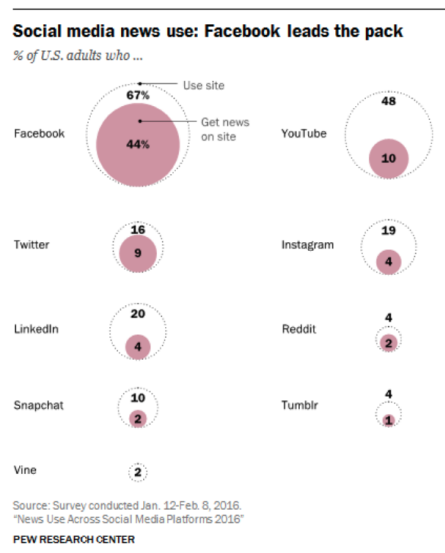


Figure B.44.: Most social media news consumers only get news on one site - 2016 [206]

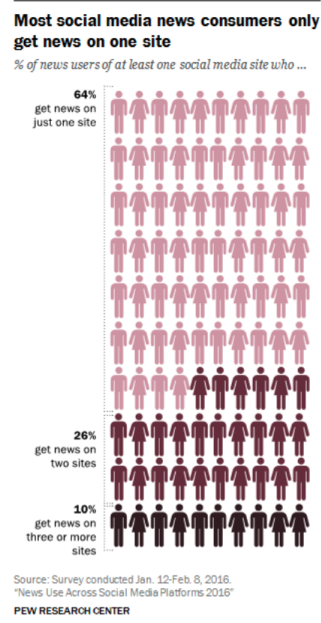


Figure B.45.: Get information by searching or by chance? - 2016 [206]

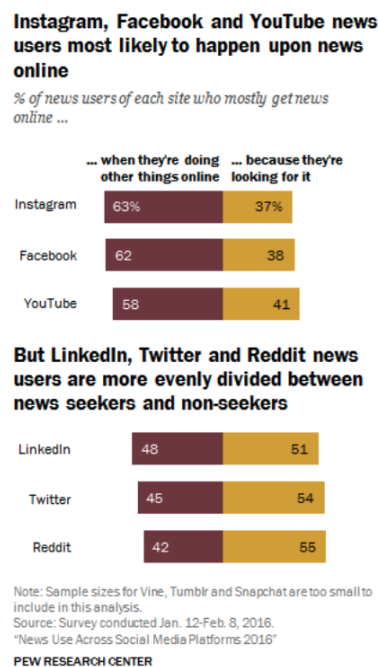


Figure B.46.: Growth in use of social media for news - 2016 [206]

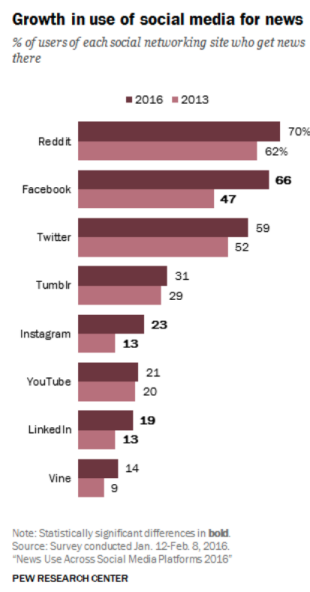


Figure B.47.: Two-thirds of U.S. adults get news from social media - 2017 [205]

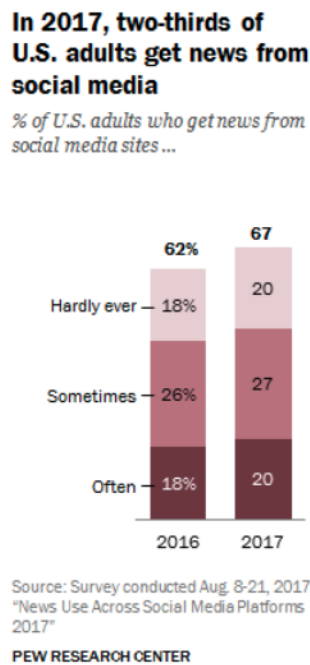


Figure B.48.: Twitter, Youtube and Snapchat grown since 2016 in portion of users who get news on each site - 2017 [205]

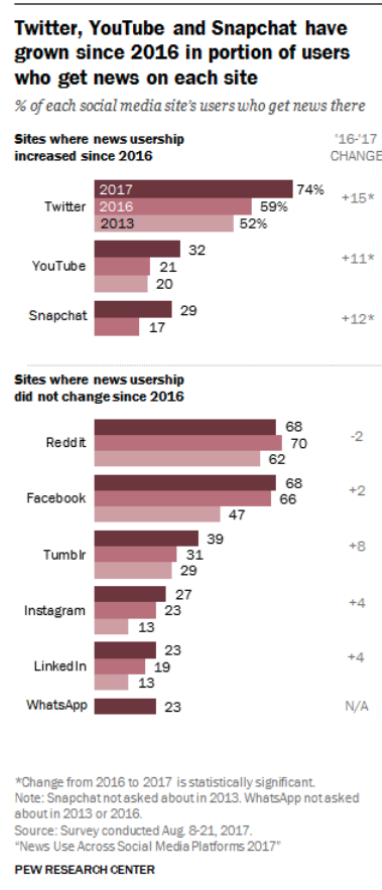


Figure B.49.: Social media sites as pathways to news - 2017 [205]

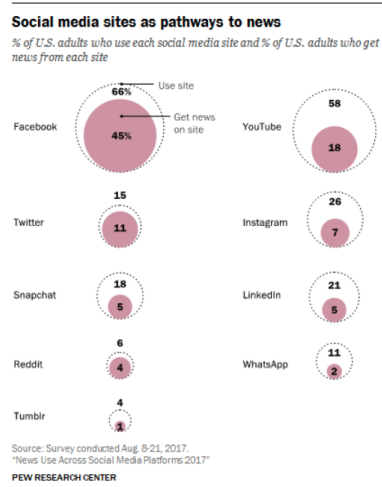


Figure B.50.: More Americans get news on multiple social media site - 2017 [205]

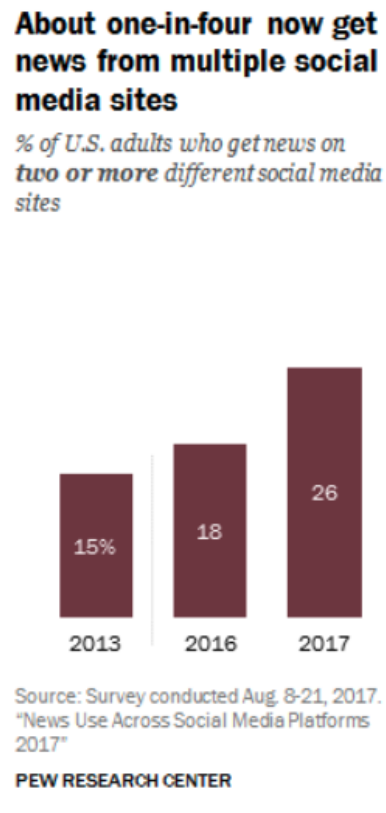
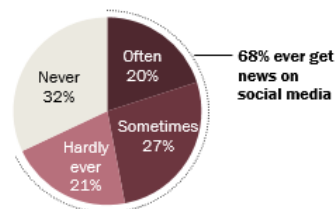


Figure B.51.: Americans get their information from social media, but a majority question its accuracy - 2018 [228]

About two-thirds of Americans get news on social media

% of U.S. adults who get news on social media ...



But most social media news consumers expect news there to be inaccurate

% of social media news consumers who say they expect the news they see on social media to be ...



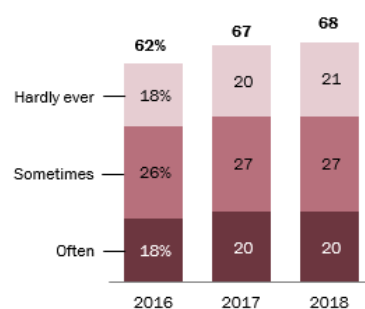
Note: No answer responses not shown.
Source: Survey conducted July 30-Aug. 12, 2018.
"News Use Across Social Media Platforms 2018"

PEW RESEARCH CENTER

Figure B.52.: Growth in social media news consumption slows down - 2018 [228]

No change in share getting news on social media in 2018

% of U.S. adults who get news on social media ...



Source: Survey conducted July 30-Aug. 12, 2018.
"News Use Across Social Media Platforms 2018"

PEW RESEARCH CENTER

Figure B.53.: Social media sites as pathways to news - 2018 [228]

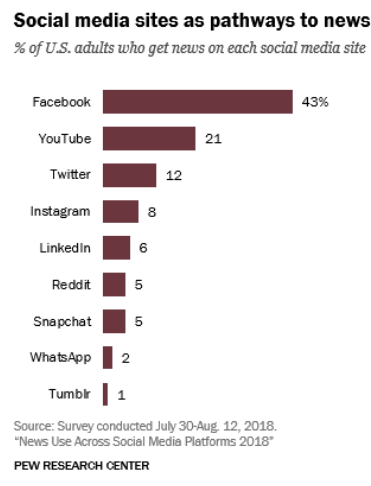


Figure B.54.: Reddit, Twitter, Facebook stand out as sites with the most news-focused users - 2018 [228]

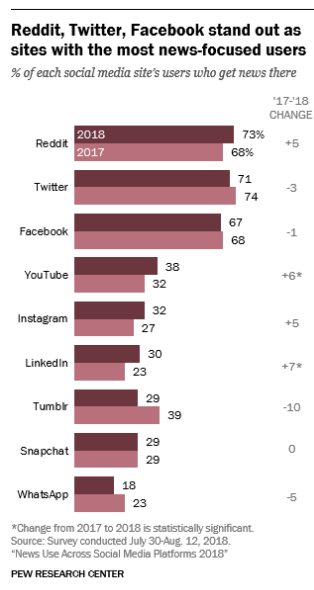


Figure B.55.: Most social media news consumers are concerned about inaccuracy - 2018 [228]

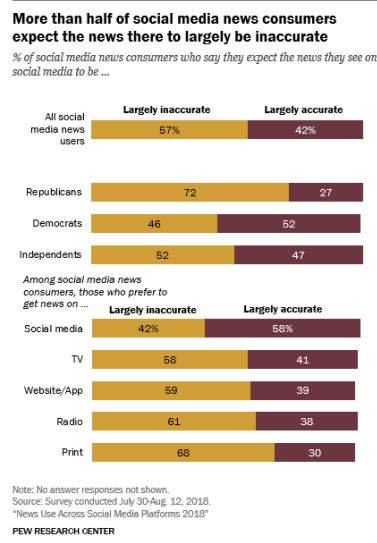


Figure B.56.: About half of Americans get news on social media at least sometimes - 2020 [70]

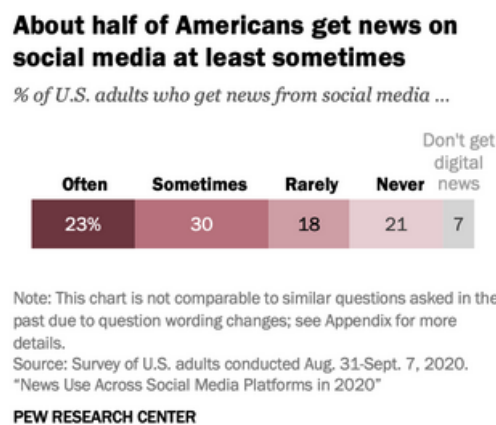


Figure B.57.: Facebook stands out as regular source of news for Americans - 2020 [70]

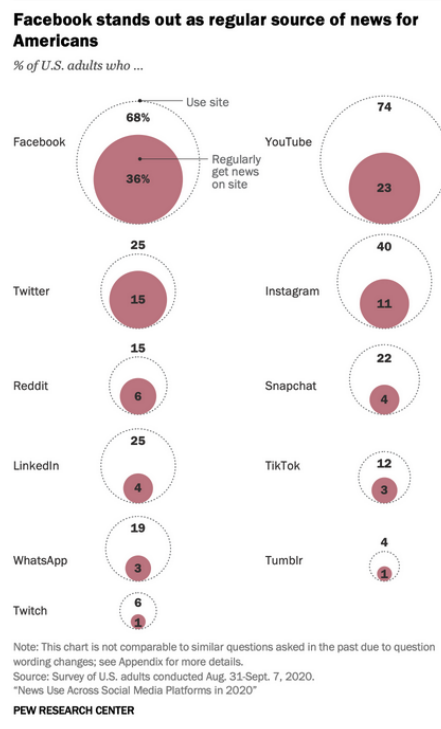


Figure B.58.: Large portion of Twitter users regularly get news on the site - 2020 [70]

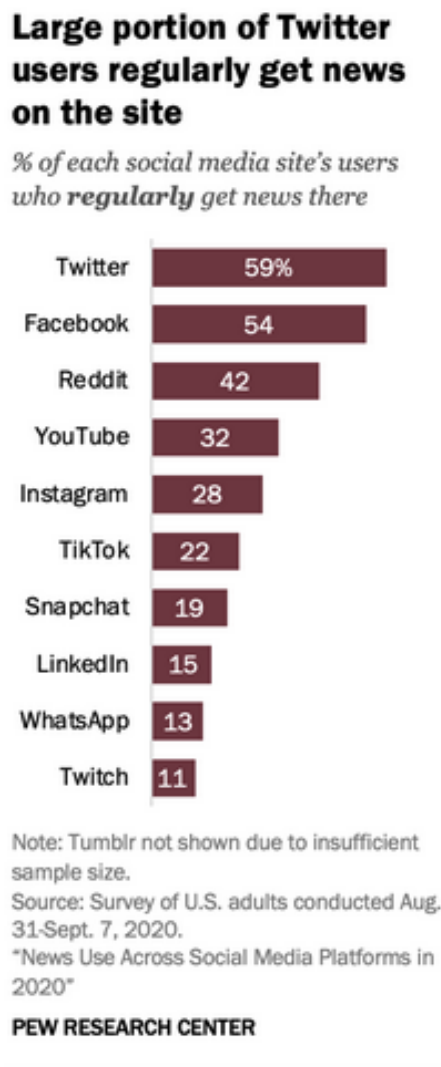


Figure B.59.: As in past years, most social media news consumers expect news there to be inaccurate - 2020 [70]

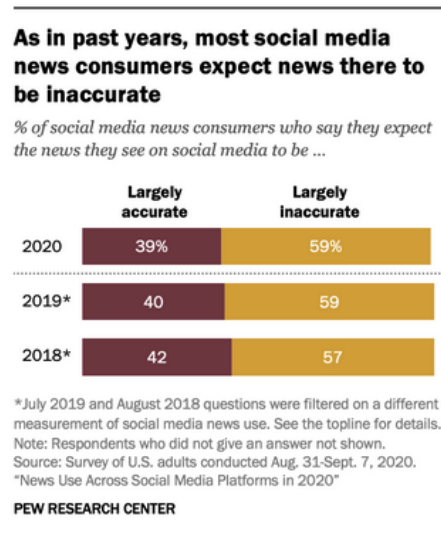


Figure B.60.: About half of Americans get news on social media at least sometimes, down slightly from 2020 - 2021 [249]

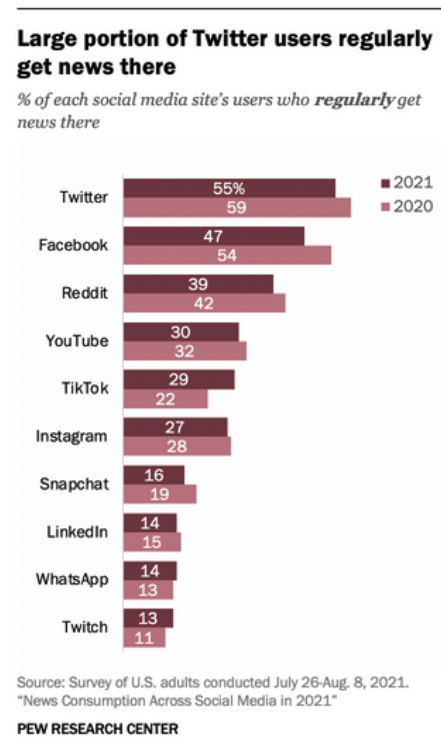


Figure B.61.: Nearly a third of Americans regularly get news on Facebook - 2021 [249]

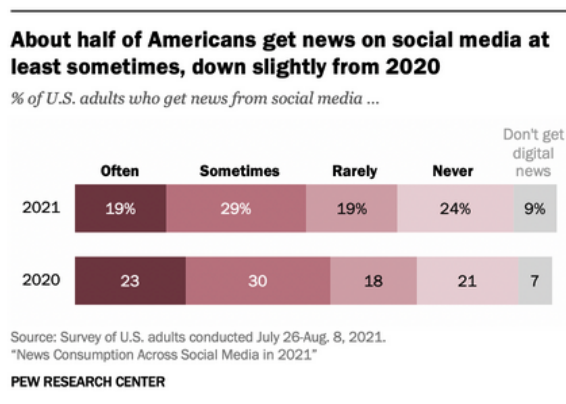
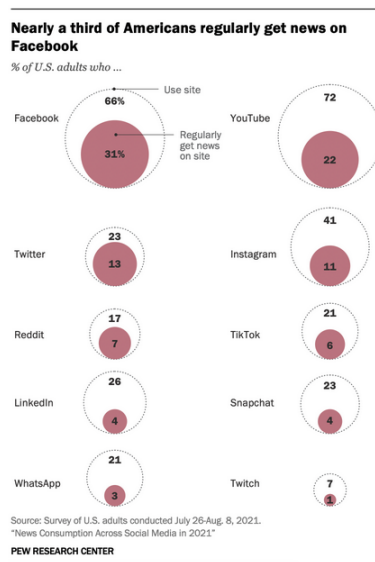


Figure B.62.: Large portion of Twitter users regularly get news there - 2021 [249]



B.9. We are impatient

The Internet has accustomed us to receiving everything instantly. It has also made us lose our ability to be patient. We're constantly looking for our dopamine, which makes us aggressive. On the Internet, this translates into the fact that we are no longer wait to receive information or what we are looking for.

Here are some figures: If a site takes more than a second to load, 40% of the users leave it

and people spend, on average, less than a second [351] or even less than 10-20 seconds on a web page [218]. 4 out of 5 users will click away if their video stalls while loading and 60% of users stop watching a video after 2 minutes [28]. Office workers check their email 30 times an hour [218]. A person looks at his phone on average 1500 times a week [218], ... In the 2015 Microsoft study, 77% of people between 18 and 24 years old answered positively when asked if the first thing they do when they have nothing to do is to take their phone [188][234]. We are used to being bombarded with information, to not having the patience to wait for a page to load, to perform many tasks at the same time, to be constantly distracted, ... it increases our aggressiveness, makes us more fragile and all this because we are looking for dopamine: we want to be bombarded with information and have it arrive instantly when we scroll.

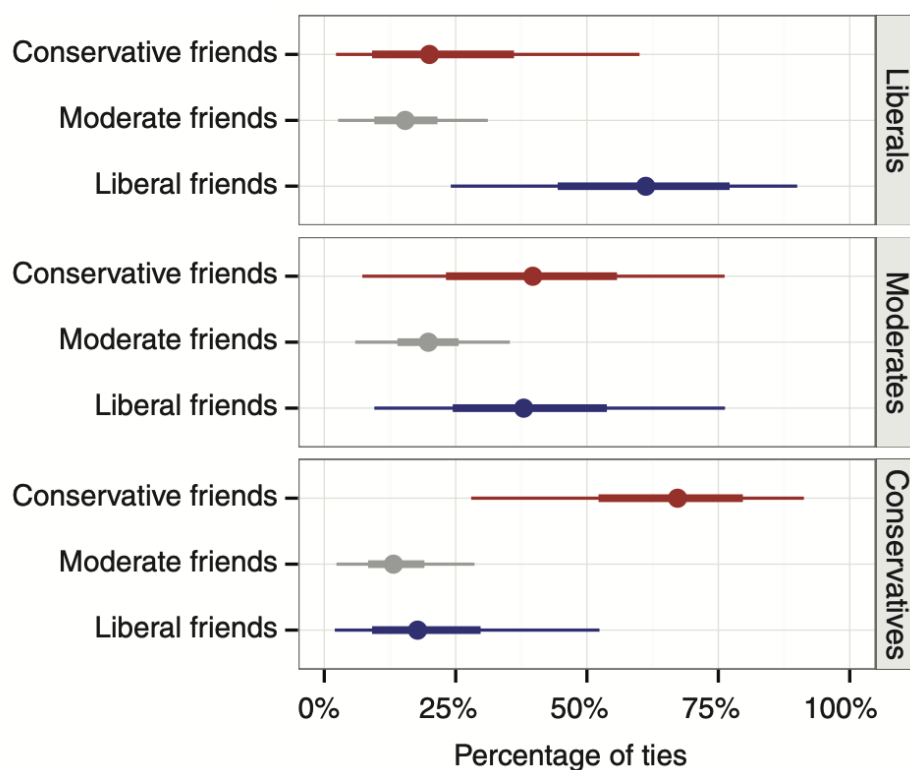
B.10. A society more divided than ever

B.10.1. Democrat and Republican, more divided than ever

Most of the time, we surround ourselves with people who have the same ideologies as us and our political orientation is mainly influenced by our family [148]. This does not mean that we are not confronted with other ideologies but we often agree on several points with people close to us. In our survey, which you can find in Appendix D.1, at the question 11/15, we see that 75% of people share the same ideas about the news on at least 50% of the topics. We notice this trend is the same on social networks. In fact, according to a study conducted in 2015 [92] based on data from 10.1 million U.S. Facebook users, they found that conservative people tend to have more conservative friends (see figure B.63). Same finding for liberal people but less pronounced and moderate people have more diversity.

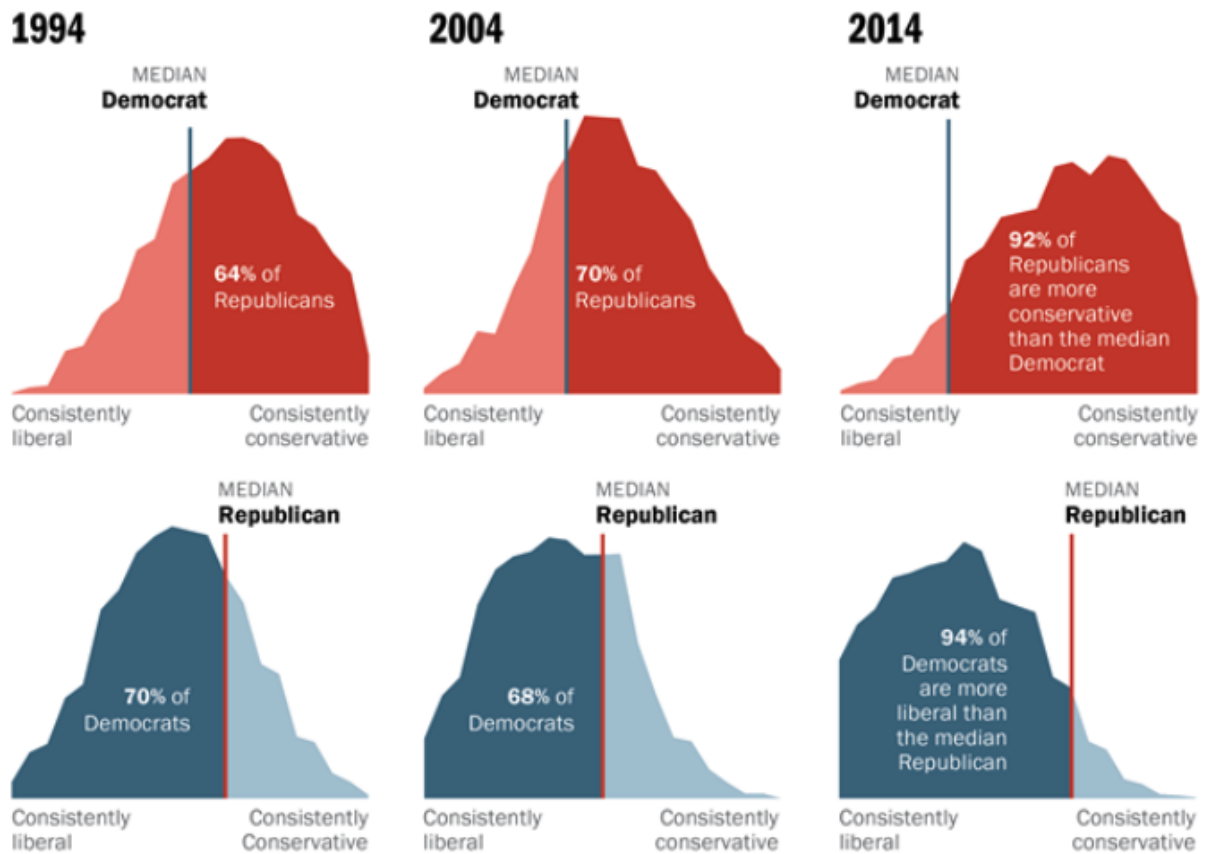
Figure B.63.: Homophily in self-reported ideological affiliation. [92]

Fig. 2. Homophily in self-reported ideological affiliation. Proportion of links to friends of different ideological affiliations for liberal, moderate, and conservative users. Points indicate medians, thick lines indicate interquartile ranges, and thin lines represent 10th to 90th percentile ranges.



Moreover, in a 2014 telephone survey of more than 10,000 Americans conducted by the Pew Research Center [255], 50% of conservatives and 35% of consistent liberals say it is important, for them, to live in a place where people share their political ideology. They also found that the Republicans are increasingly to the right and vice versa for the Democrats (figure B.64).

Figure B.64.: Republicans Shift to the Right, Democrats to the Left. [255]

Republicans Shift to the Right, Democrats to the Left*Distribution of Republicans and Democrats on a 10-item scale of political values*

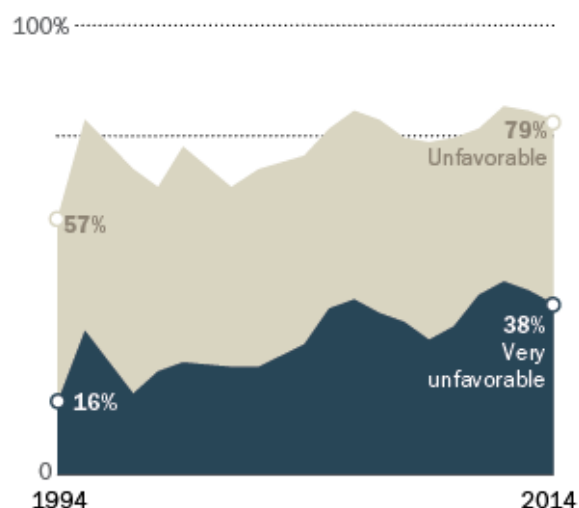
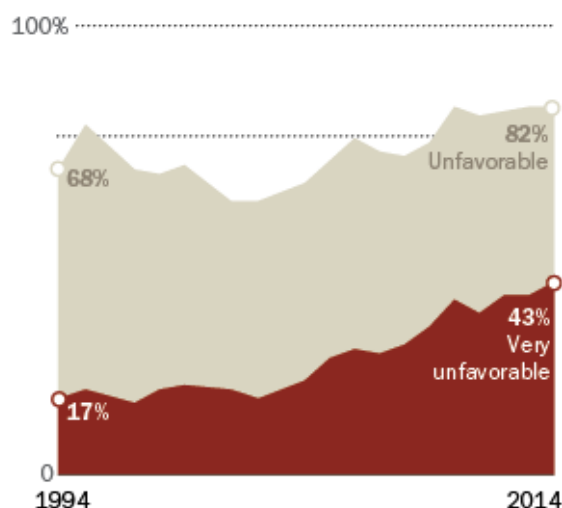
Source: 2014 Political Polarization in the American Public

Notes: Ideological consistency based on a scale of 10 political values questions (see Appendix A). Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents (see Appendix B).

PEW RESEARCH CENTER

They have noticed that there is more and more hatred between opposite parties (see figure B.65).

Figure B.65.: Growing Partisan Antipathy. [255]

*A Rising Tide of Mutual Antipathy***Democratic attitudes about the Republican Party****Republican attitudes about the Democratic Party**

Source: 2014 Political Polarization in the American Public

Notes: Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents.

PEW RESEARCH CENTER

B.10.2. What kind of world do we live in?

We live, now, in a world built around technological companies like Meta, Google, Microsoft, Apple, Amazon, ... Our consumption and production methods have become dependent on them. Now they dictate how we should behave, how we should consume, how we should work, etc. It has become inconceivable to live in a world without computers, without Internet, without instant messaging, without connection, without everything else we can have instantly and for free. Now, the people who will be put forward are not the ones who make new discoveries, who contribute to improve the society, who find solutions to our current problems (political, ecological, economic, ...), etc. No, they are the ones who manage to make sensationalism by any means at the expense of quality content and who contribute to the development of the society. All recommendation algorithms will prefer to show you content that will capture your attention, create interaction and therefore generate more money. And that, the creators of content (Youtubers name themselves like this) have understood it well. They have access to

tools that allow them to track how people interact with their video. Like knowing when they left the video. Youtubers struggle to understand the algorithm and create content where the videos are edited in a super dynamic way so that we get bored as little as possible, everything must go fast, etc. The video thumbnails are designed to make us want to click on the video, etc. Everything is calculated. It's not the fault of the Youtubers, it's Youtube itself that imposes these standards otherwise you are not put forward[275]. We are in a system where the rich get richer (the more you are seen, the more you will be seen).

For average people, the same thing is true but in a different context. People are looking for popularity, social validation, who will have the most likes and followers. Some people judge each other based on the number of likes and followers they have. Social networks have become a tool for measuring their relative popularity. [170]

Here are 3 examples (we limit ourselves to 3 because there is an infinite number of examples), which in our opinion, illustrate well the system in which we live:

Firstly, we see since a few years the omnipresence of influencers on social networks who have a lot of followers and generate a lot of money with product placements that are often scams. There is also the case of an influencer who was selling her bath water. This has made so much talk that she was invited in the show of Cyril Hanouna: TPMP. We see the level of what is put forward by social networks and media to have an audience...

Secondly, do you know the first Youtubeur in the world? It's MrBeast. He has 150 million subscribers and makes between 100 and 200 million views (on average) per video at the moment (2 March 2023). To be number one in the world, we wonder what wonderful content he offers us? Well, here are some titles of these videos: A \$1 plane ticket vs. \$500,000, I offered a private island to my 100,000,000th subscriber, If you can carry \$1,000,000, you keep it, The last one to leave the bathroom wins \$1,000,000 (Part 3) or Hydraulic press VS Lamborghini. The world number one has videos that make other people win phenomenal amounts of money with stupid concepts like the last one who takes his hand out of the car wins the car. We can see that we are on sensationalism with videos that are not very complicated to set up with stupid concepts which do not require intellectual effort and where the only objective is to consume a maximum, pollute and make believe to spend \$1 000 000 in a car and then explode it is cool. Feel free to watch this video [187] (from the Youtuber 'HugeDécrypte') which shows who MrBeast is and from which we can draw conclusions that are not specifically pointed out by the video.

Third and last example is the Youtuber Logan Paul. To put it in context, he notably made a polemic when he posted a video on Youtube where he filmed dead people (they had committed suicide) in a forest in Japan. Now he makes, in particular, podcasts with replay on Youtube (channel: IMPAULSIVE) where he interviews famous people. One of the interesting interviews is the one with Dana White where Dana White laughs and says *where did covid go?* implying that it never existed. It's interesting to see that because when you look at the comments, a lot of people admire this person and agree with him. At the same time, we can see a lot of fake news spread on the networks as for example with the covid and that creation of anti-vax groups.

The problem with these Youtubers is that they have a rather young and very manipulable audience. They produce low quality content that attracts attention by making millions of views.

B.11. Study of the mentality of companies like Meta

This section is mainly based on the Netflix documentary The Social Dilemma released on September 9th 2020 [264]. This documentary is one of our sources of motivation for this project. It highlights several issues related to social networks that we will go through in this section. In the documentary, they had the opportunity to interview some big names in the field of social networks that we will list in the following subsection because we will, afterwards, reference who said what sentence in the next subsections. Much of the information in this section is a **transcript** from this documentary. So much of the credit goes to the Netflix documentary. But we have added some information from our own research, either to support the ideas put forth or to add other information that was not in the documentary but that we thought was interesting.

B.11.1. Names of the speakers in the Netflix documentary The social dilemma

A . Anna Lembke.

Al . Alex Roetter.

Az . **Aza Raskin.**

B . **Bailey Richardson.**

C . **Cathy O’Neil.**

Ch . **Chamath Palihapitiya.**

Cy . **Cynthia M. Wong.**

F . **Fernando Randima.**

G . **Guillaume Chaslot.**

J . **Jaron Lanier.**

Je . **Jeff Seibert.**

Jo . **Joe Toscano.**

Jon . **Jonathan Haidt.**

Ju . **Justin Rosenstein.**

M . **Marco Rubio.**

R . **Rashida Richardson.**

Re . **Renee DiResta.**

Ro . **Roger McNamee.**

S . **Shoshana Zuboff.**

Sa . **Sandy Parakilas.**

T . **Tristan Harris.**

Ti . **Tim Kendall.**

B.11.2. Topics covered in the documentary

Here are the different topics that cover the documentary. I will transcribe interventions of people mentioned in the previous point while completing from time to time the information with personal research.

B.11.3. Are we the product ?

We often hear that if you don't pay for the product, you are the product. When it comes to social networks, the real customers are the advertisers. They want to keep us as much as possible on the platform to collect a maximum of data on us. The more data they have, the more accurate their prediction model on our behavior and interests will be. [Tristan Harris][Tim Kendall][Justin Rosenstein] This is what we call surveillance capitalism (I'll come back to this in the next point). Jaron Lanier says that there is a subtlety and that's more complex than that. For him, it's the gradual, slight, imperceptible change in your behavior and perception that is the product. And that this is their only way to make money : change the way we act, the way we think, who we are and this change is gradual. This change has for goal that people go in the direction that Facebook wishes that people go. It is thus manipulation and that is worth a lot of money. So if we summarize this statement, for Jaron Lanier, it is not us personally who are the products but our behavior, our way of thinking, ... which is put on sale to their advertiser.

This is therefore directly linked to surveillance capitalism, as Tristan Harris and Shoshana Zuboff say.

B.11.4. Surveillance capitalism

For Shoshana Zuboff, social networks sell certainty: having the guarantee that when you publish something, it will be successful. To do this, you need to have the most accurate predictions possible and therefore collect a lot of data. The more Facebook announcers make money, the more Facebook will make money because advertisers will continue to invest in advertising. This market between Facebook and its advertisers is made exclusively of human data. To collect all this data, you need tracking tools.

B.11.5. Tracking

For Jeff Seibert, all our actions are tracked and measured. They know how long we spend on each post/video/photo/..., all our actions are recorded (scroll, clicks, ect), ... Everything is carefully studied and recorded. They know when we look at an image and for how long. They

know when we feel lonely, depressed, when we look at pictures of our ex-girlfriend, they know everything we do. Nothing escapes them. They can determine if you are an introvert or an extrovert, what are your neuroses, your personality traits.

All this data they collect is put into systems that have almost no human supervision and are able to make increasingly accurate predictions about what we will do and who we are [Sandy Parakilas].

All the clicks, videos we watched, ... everything is listed in their database to improve their model. With this model, they can predict our actions, our emotions to which we are most receptive, the videos we like, ... [Tristan Harris].

B.11.6. Do they sell our data to third parties ?

It is not in their interest to sell our data because they make models that try to predict with the most accuracy our character traits and our future actions. Whoever has the best model wins. So it is not in their interest to sell this data [Aza Raskin]. It would allow other competitors to make better models.

For Tristan Harris, companies like Meta have 3 mains goals :

1. Commitment: make sure we stay connected as long as possible.
2. Growth: they have to make sure that we come back and that we encourage as many of our friends as possible to come to the platform.
3. Advertising: to generate as much money as possible.

And all this is managed by algo's whose mission is to find what to show us.

Tim Kendall gives us an example of what could happen during a morning meeting with Mark Zuckerberg. He explains that at Facebook, they often thought about what parameters to adjust to boost ad revenue and that it was easy. Also, Mark Zuckerberg could sometimes come in and say *I want Korean users today* and boom, they would surgically adjust the ads (they know how to be very precise).

B.11.7. A world where online communication prevails over all others

For Jaron Lanier, we live in a world where online communication prevails over all others and in this connected world, every time 2 people exchange, there is a 3rd actor who interferes in the conversation to manipulate them and maximize his profits.

Always according to Jaron Lanier, we have created a generation that has been immersed in this system from a very young age. Each of their online communication and interaction is done exclusively through manipulation. Deception and deviousness are at the very heart of online communication.

B.11.8. Magic and social networks

Any sufficiently advanced technology is indistinguishable from magic. Arthur C. Clarke

Magic allows you to trick and deceive people. With social networks you can do it too.

The magicians were the first neuroscientists of their time. They were the first to understand how the human brain works. They managed to detect and understand a part of our mind that we do not control. This is what makes the illusion work. Our mind is vulnerable no matter who we are or how much we study [Tristan Harris]. Tristan Harris studied at Stanford and explains that he learned how to translate human persuasion techniques into technology. So there are courses dedicated to that. Sandy Parakilas adds that there are many people from silicon valley who went through these courses (people from Facebook, Uber, Linkedin, ...). They learned how to make technology more persuasive and Tristan Harris is one of those people.

B.11.9. Persuasive technologies (manipulation and addiction)

Persuasive technologies are technologies pushed to the extreme with the sole purpose of changing the behavior of a person. The goal is that people do what you want them to do, like scrolling endlessly on a screen [Tristan Harris]. At Netflix, Snapchat and Facebook, specialized engi-

neers, the Growth Hackers⁵, have a single mission: to design new ways to encourage us to spend more time on their platforms. These so-called persuasive technologies are mostly based on the theory of psychologist B. J. Fogg, taught at Stanford[149].

Each time we refresh the page, the results change on Facebook. In psychology, this is called intermittent positive reinforcement [Joe Toscano]. The intermittent positive reinforcement is the procedure by which the probability of occurrence of a behavior tends to increase following the addition of a contingent appetitive stimulus to the response. For example : addition of a reward, congratulations... [8] So, with social networks, when you refresh the page, you are rewarded with new content. The fact of scrolling, refresh the page, ... gives us rewards and provides us with dopamine (which makes us addicted) [314] and so we always want to get our dose of dopamine and we repeat these gestures instinctively. With time, we are not aware of what we are doing. We do all these actions in an automatic way.

Tristan Harris explains also that when we look at our phone, we don't know if we are going to receive something or even what we are going to receive as a notification. It's the same principle as the pay machines. The fact that we use the service is not enough for them, they want to get deep into our brain and our subconscious, that we are an unconscious habit that makes us conditioned to do a task without even realizing it. Every time we look at our cell phone, it might be the jackpot with a notification. It's not random, it was designed that way. Tristan explains that personally, it's true that he sometimes open Facebook without even realizing it when he's on his phone. It's like an instinct. I've already had the experience of closing Facebook on my phone and then reopening the application directly. So there are things I do instinctively.

Another example is the photo tag. We receive a notification/email that we've been tagged, of course we'll open it to look at the photo (it's human, we're intrigued). The question is to know why the email/notification does not contain directly the photo [Jeff Seibert]. Tristan Harris explains also that when Facebook discovered this feature (to make notifications without putting the content), they did it thoroughly because they realized that people were going to do that all day because it intrigues them to know what they were tagged on.

⁵Growth Hacking is a set of techniques, or rather a state of mind, a methodology, whose objective is the rapid growth, by all means, of a service or a product [334]. A hacker is a person who uses unconventional means to access a system. The growth-hacker uses unconventional means to grow a user base. He can work on all stages of the customer acquisition process (acquisition - conversion - retention) to achieve his growth objectives [291]

B.11.10. A department dedicated to persuasive technology

There are teams entirely dedicated to develop persuasive technology, to activate the economic growth of society. They hire engineers whose only mission is to get into people's heads to get them hooked. The goal is to increase growth, to have more and more people connected, ... [Tristan Harris]. They are therefore Growth Hackers.

Chamath Palihapitiya, former Vice President of Audience Growth at Facebook explains that their goal when he was at Facebook was for each new member to have at least 10 new friends in 7 days.

Sandy Parakilas explains that Chamath Palihapitiya used techniques where he did several tests on the Facebook platform and watched how users reacted. With all these tests they were able to determine in an optimal way what the users expected from their service. So it's manipulation. We're like lab rats except that here, there's no benefit for us, they're only doing this to make money by manipulating us

Shoshana Zuboff explains that there has been an experiment that was done at Facebook that they called large-scale contagion experiment. The idea was to introduce subliminal messages on Facebook's pages to convince more people to go vote in the midterm elections (So it's manipulation). It worked perfectly, there are more people that voted. So they know the power they have and what they are capable of. They are able to affect real life without the user ever really being aware of what has happened. They are also able to play with our emotions, manipulate us, ...

To give you more details about this experiment [96] [115], it was conducted during the 2010 US congressional elections on 61 million Facebook users. The question, at the beginning, was whether online networks, which harness social information from face-to-face networks, can be used effectively to increase the likelihood of behaviour change and social contagion. So they did an experiment with politics as voting turnout is significantly correlated among friends, family members and co-workers in observational studies. So they separated these 61 million people into 3 groups randomly :

1. **A social message group.** They ($n = 60\,055\,176$) were shown a statement at the top of their News Feed. This message encouraged the user to vote, provided a link to find local polling places, showed a clickable button reading I Voted, showed a counter indicating

how many other Facebook users had previously reported voting, and displayed up to six small randomly selected profile pictures of the users Facebook friends who had already clicked the I Voted button.

2. **An informational message group.** They ($n = 611\,044$) were shown the message, poll information, counter and button, but they were not shown any faces of friends.
3. **A control group.** They ($n = 613\,096$) did not receive any message at the top of their News Feed.

They do a lot of rigorous and precise analysis in the report that I'm not going to go into here but I will give you the important points of their analysis. They put forward two very important things.

The first important thing they put forward is the fact that it is not the messages that have been put on people's feeds (direct effect) that have manipulated people the most but the fact that it has generated more communication between people on this subject (indirect effect). Let me explain, they measured that out of 61 million users, they have an average of 149 friends on the platform. On this 149 friends, they have more or less strong connections (acquaintances or close friends). This is what we call weak ties and strong ties. Other research indicates that close friends have a stronger behavioral effect on each other than do acquaintances or strangers. We therefore expected mobilization to spread more effectively online through 'strong ties'. They measured the number of strong ties each person had (they explain in detail how they measured this in the report) and found that each person had at least one strong tie and on average they had 10 strong ties. They had interesting results. They realized that whether or not our close friends saw the message had an influence on whether we would vote or not. Thus, they conclude that ordinary Facebook friends (= weak ties) may affect online expressive behaviour, but they do not seem to affect private or real-world political behaviours but close friends (= strong ties) seem to have influenced all three. It should be noted that the figures he gives are quite low, for example : the results suggest that users were about 0.011% more likely to engage in an act of political self-expression by clicking on the I Voted button than they would have been had their friend seen no message. So, as they explain, the magnitude of these contagion effects are small per friend if we look at the figure, but it is important to remember that they result from a single message, and in many cases it was not possible to change the targets behaviour.

Second, they measured the number of votes implied by this manipulation. There was 60 000 voters due to direct effect of the message on the feed and 280 000 voters due to indirect effect (social contagion through strong ties). The total is therefore 340 000 votes which represent 0,14% of the total number of voters (236 millions in 2010).

They therefore conclude that the efforts to influence behaviour should pay close attention not only to the effect a message will have on those who receive it but also to the likelihood that the message and the behaviour it spurs will spread from person to person through the social network and that online mobilization works because it primarily spreads through strong-tie networks that probably exist offline but have an online representation.

We must realize that the power of these platforms does not come exclusively from the fact that they can choose what they want to show/hide us, but from the networks they have to spread information. Moreover, depending on where they are going to put the information in their network, it will not propagate in the same way (depending on the weak ties and strong ties) and they know very well where they have to inject the information in their network so that it reaches the people they want. They know how to control the propagation of an information because they have all the information about their network (the connections between their users, the interests of each one, ...) and where is each information in it. It's really like if they could propagate a virus while controlling at 100% the contamination rate and who must be contaminated. Moreover, we see that the fact of propagating an information in their networks also has an offline incidence because the people will speak about it when they see themselves. They can therefore influence the conversation of people outside the platform. The fact of having used a counter that gave the number of people who voted can give a Rich-Get-Richer effect. Indeed, the fact to see that there are many people who vote (and especially friends), it influences us to also vote. So we can see that the way the information is represented can also have an effect on the strength of the manipulation.

To return to the Netflix documentary, Chamath Palihapitiya adds that the goal of these companies is to manipulate us psychologically as much as possible with a dose of dopamine as a reward (= positive reinforcement as explained before). We find this process at Facebook, Instagram, WhatsApp, Snapchat, Twitter, ...

So in the end, a vulnerability is exploited and in this case, it's the human mind. Moreover, they (Sean Parker, Mark Zuckerberg and Kevin Systrom) understood that (that they were exploiting a vulnerability in human psychology) consciously but they did it anyway [Sean

Parker] [62]. Social networks are not tools that wait to be used, they have their own objectives and to achieve them, they use psychological techniques against us [Tristan Harris]. Tristan adds that they use AI against us to try to understand how the human brain works in order to discover what triggers an emotional response.

B.11.11. Is social networking a drug ?

Tim Kendall was the former president of Pinterest. He explains that he himself couldn't get away from his laptop when he got home to take care of his kids. He was either writing emails or scrolling through Pinterest. He found it ironic that during the day he was working on this very addictive project (Pinterest) and at night he was falling into his own trap. He knew what was going on but he still fell into the trap. He couldn't control it (everyone else working on these projects says this too, even though they know how it's all designed, they still remain receptive).

Moreover, as explained earlier, these social networks are designed to give us doses of dopamine (= reward). As Sean Parker said : The thought process that went into building these applications, Facebook being the first of them was all about: How do we consume as much of your time and conscious attention as possible? That means that we need to sort of give you a little dopamine hit every once in a while, because someone liked or commented on a photo or a post or whatever. And that's going to get you to contribute more content and that's going to get you more likes and comments[62].

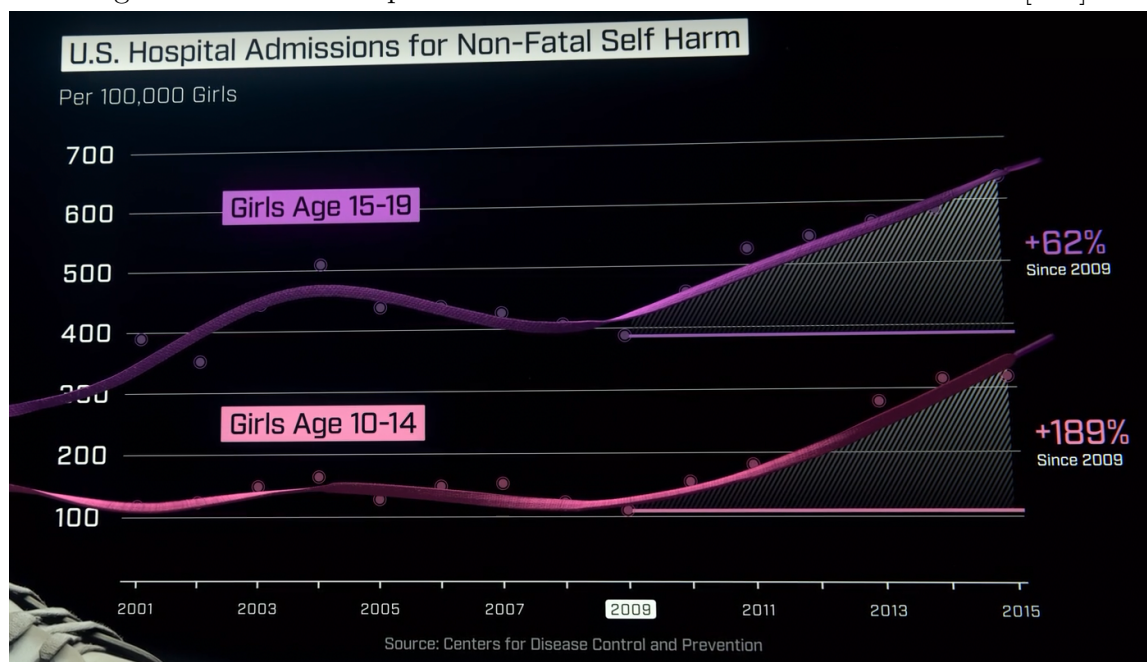
Dr. Anna Lembke explains that social networks are drugs because humans constantly need to feel connected to other people. This releases a dose of dopamine and affects the reward circuit in the brain. The basic purpose of this system that our body has put in place is that we come together, that we live in community in order to reproduce and that our species endures. So there is no doubt that tools like social networks that allow to make this connection between people can potentially make people addicted.

All these algorithms have not been designed by careful people but by people who want these algorithms to recommend us what corresponds best to us or to subliminate us with all the filters that we can have for our photos. But that's not all, they also allow to control our attention because these algorithms are very powerful to get into the brain of children and modify their self-esteem, their identity. We have evolved to care about what other members of our tribe

think of us. It was and is important that we know what others think of us, but have we really evolved to care about what 10,000 people think of us? We weren't designed to be bombarded every 5 minutes with social validations [Tristan Harris]. We build therefor our lives around this false sense of perfection. We are rewarded in the long run with hearts, likes, ... we confuse this with real value and truth. It is a false and brutal popularity. How does it feel to get a lot of likes? The truth is that we feel empty inside. This makes us enter a vicious circle where we wonder *what's next, I want this feeling again*. Let's multiply this feeling by 2 billion people and think about how people feel about each other. It has become very serious [Chamath Palihapitiya].

To give you some graphs to illustrate that ⁶, Jonathan Haidt explains first that there is an increase in depression and anxiety among young people and this growth began between 2011 and 2013. It is important to know that Facebook is created in 2004 but remains within the university campuses. Then, in September 2006, it becomes open to the general public. Finally, it is in 2009 that social networks become available on smartphones.

Figure B.66.: U.S. hospitalisation admission for non-fatal self harm. [264]

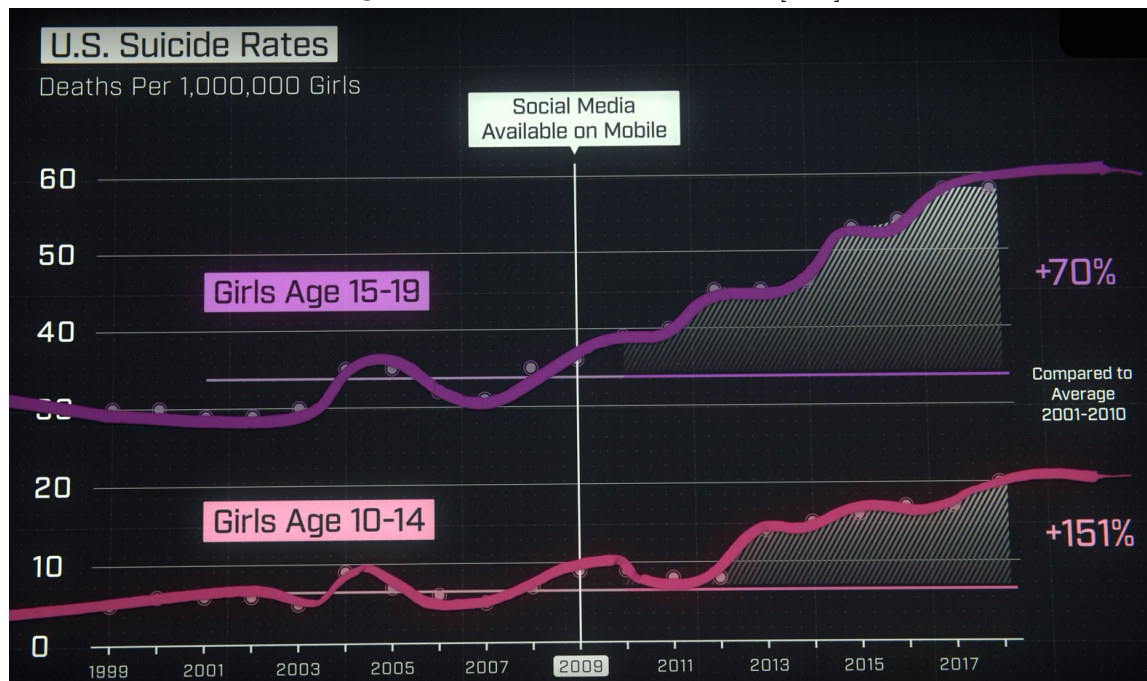


Then, he explains that the same is true for the suicide rate for young people (see figure B.67)

⁶Those graphs come from the documentary Netflix where, for the figures B.66 and B.67, they get the figures from the Centers for Disease Control and prevention. For the figure B.68, the source come from NVIDIA.

and Tim Kendall says that those service kill people, people commit suicide because of it.

Figure B.67.: U.S. suicide rates. [264]



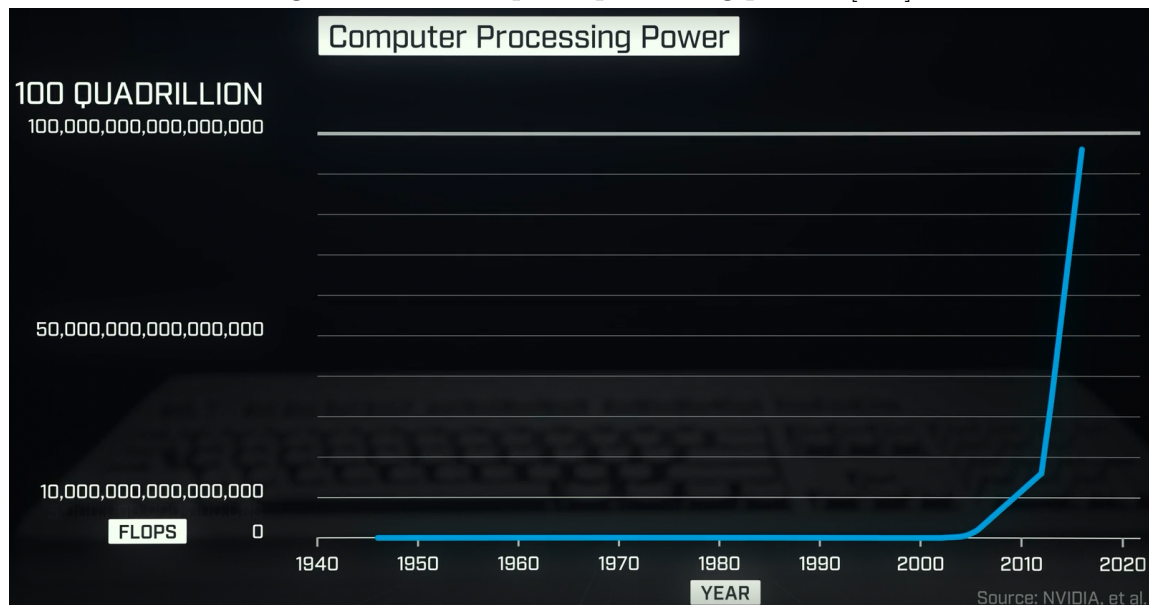
There is a correlation between the appearance of social networks on smartphones and therefore accessible wherever you are if you have a connection (wifi, 3G, 4G, ...) and the increase in the rate of suicide, depression, anxiety among youth. Jonathan Haidt explains also that the Z generation (born after 1996) is the first generation to have had access to the networks as soon as they arrived at college. When they came back from college, they were directly on their phone. It is a more anxious and fragile generation, which is more afraid to take risks. For example, we note that they pass their license later and later. We also note that for the love relationship: the dating rate is down. So finally, social networks that are supposed to have the goal of bringing us closer, separate us more than anything else.

Tristan Harris adds that the conditioning of people is done from a very young age now. No one wants social networks to manipulate our children, to prevent them from doing their duty. Before there was a time slot where there was no advertising on TV for cartoons. Now, with YouTube kids and all these platforms, they are exposed to it like everyone else. All the past regulations are gone. We are conditioning and training a whole generation of people. Now, when we feel uncomfortable or lonely, when we have doubts or are afraid, we have this little

object in our pocket to comfort us. The problem is that this atrophies our ability to manage our emotions. We often hear that we will learn to live with these technologies, these tools. But here, we are facing something much too dangerous.

Randima Fernando explains that the most dangerous aspect is the fact that all this is driven by a technology that evolves exponentially as it is show in figure B.68. It is one of the only technologies that evolves at this speed. The second tool that evolves rapidly are the cars that go twice as fast but for the rest, the evolution is almost negligible. The worst is that our brain does not evolve as fast and has not evolved at all compared to technology. Tristan Harris adds that we have our brain which is millions of years old and in front of it, we have a screen and on the other side of the screen, we have thousands of engineers with super computers who don't have the same objective as us. So who is going to win at this little game ? We often imagine AI as robots like in the movies Terminator or I, Robot but what people don't realize is that it is already AI that leads us, the world is under its control.

Figure B.68.: Computer processing power. [264]



B.11.12. IA and social networks

According to Cathy O'neil, algorithms are opinions in the form of code. So they are not objective because they are optimized to meet a certain definition on success. So those of

Facebook are in their economic interest (that's their definition of success). Jeff Seibert also explains that we give the algorithm our goals and the computer will learn how to achieve it autonomously. Every day, it improves by choosing the right order to show each element and nobody understands exactly how it is done because it is the computer that chooses and it is the only one to understand what it does. The algorithm has its own way of thinking and even for the one who wrote it, it can be indecipherable. You build the machine and the machine ends up building itself [Bailey Richardson]. For Sandy Parakilas, it is consequently a very complex system. At Facebook, Twitter, ... only a handful of people understand how it works and even they don't always know what can happen in this or that situation! As humans, we have almost lost control of these systems. They control the information we see because they have learned to show us what will keep us as much as possible on the platform without humans necessarily understanding why the computer has made these choices. Therefore, computers and algorithms control us ! To finish this subsection, Roger McNamee adds that when we are on Facebook, we are fighting against an AI that knows absolutely everything about us, it can anticipate all our movements. While we don't know anything about it except that we see chat videos and birthdays. It's a totally unfair fight.

B.11.13. The problem of news feeds on sites like Facebook

Jaron Lanier introduces this subsection with an example. If we take the case of Wikipedia, the content of the same page does not change depending on the person who comes on the platform. So it's one of the only things online that we all have in common. Imagine if Wikipedia said one day that they were going to give us all different content for each new visitor and people would pay us for it. Wikipedia would start spying on us, calculating what they need to do to get our attention. For the sake of profit, they would change their content to maximize their advertising revenue. This is what happens on Facebook, YouTube, ...

Justin Rosenstein gives another example: if we type in google climate change is, we will have different results depending on our location in the world. In some cities, we will get *climate change is a scoundrel*, in another city *climate change is destroying nature*. It is not the truth that is put forward when we do our research but the result is according to the place where we are and the interests that google knows about us. So I tested this myself and the results confirm that. I typed *The war is* (in French) on my browser and I was located in Belgium (so

I turn off my VPN), the figure B.69 show you the proposition of Firefox. I tested the same (on the same computer) but I turn on my VPN to be located in Ukraine, the figure B.70 show you the proposition. Last test, I was on another computer, with my VPN turn on to be located in Ukraine and with the browser Edge, the figure B.71 show you the result. First thing, if we compare the difference between figure B.69 and B.70 where I have only changed location, we realize that there are similar proposals but not in the same order. Another important thing is the sentence I framed in yellow in both figures. It is the same sentence but with a subtlety on one word. In Belgium, the sentence is *war is a too bad thing to be entrusted to the military* and in Ukraine, the sentence is *war is a too serious thing to be entrusted to the military*. So we see that there are similar proposals but with a subtlety on some words. If we look at the proposals that the browser Edge has given me (figure B.71), we can see that the proposals are completely different. What's confusing is that on the Edge browser, it's more about *the war is over* whereas with the Firefox browser, the proposals are about cunning, *necessary to have peace*, ... so it's a completely different register. Maybe it's not only due to the browser but also due to the fact that I changed my computer to use Edge and therefore the profile they created about me (my interests, ...) on both computers is different.

Figure B.69.: Proposition of Firefox for the sentence "the war is" (browser : Firefox, location : Belgium)

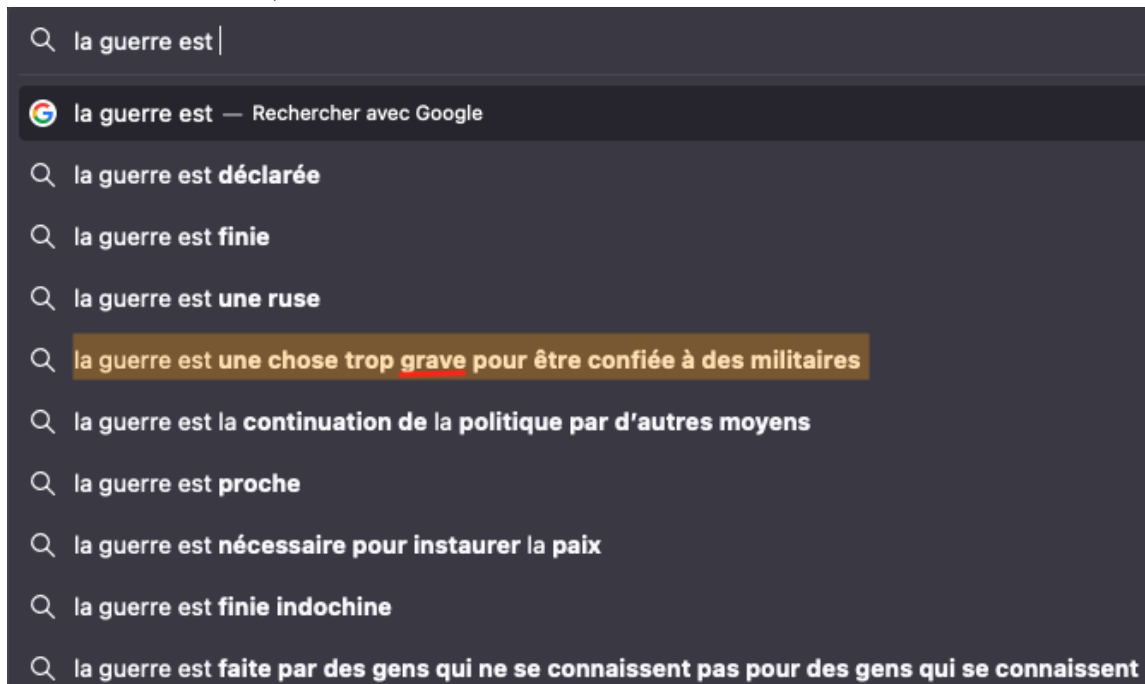


Figure B.70.: Proposition of Firefox for the sentence "the war is" (browser : Firefox, location : Ukraine)

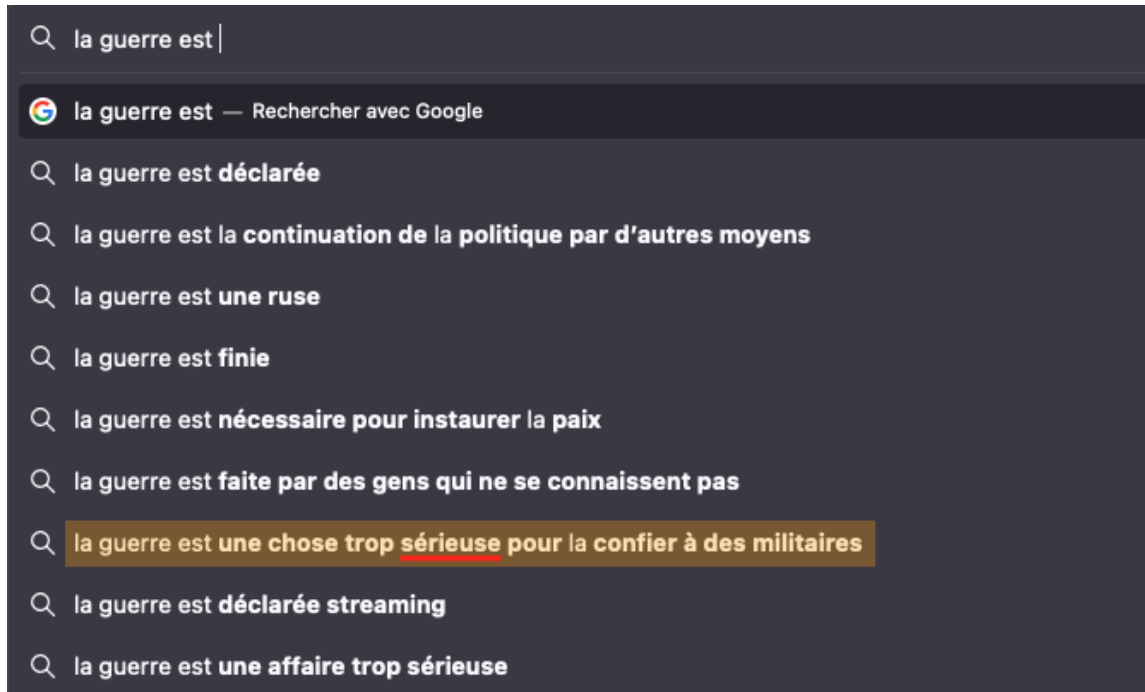
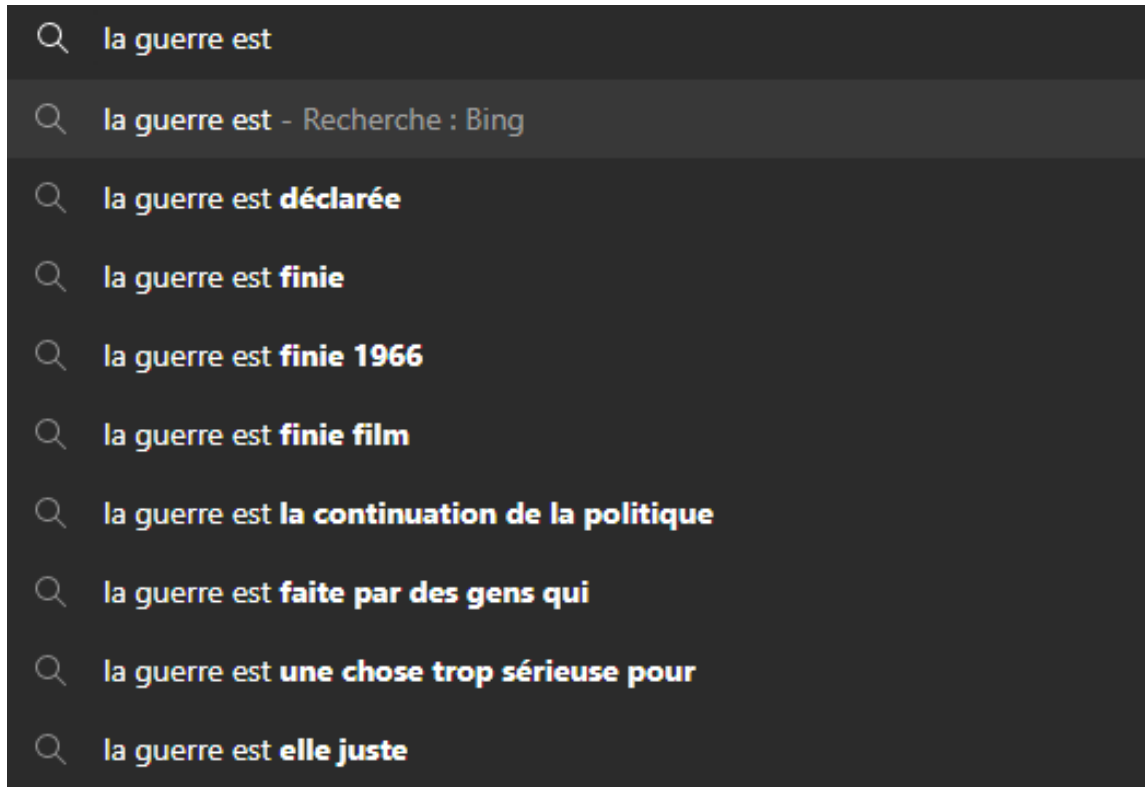


Figure B.71.: Proposition of Edge for the sentence "the war is" (browser : Edge, location : Ukraine)



To return to the subject of the news feed, Tristan Harris explains that if we take two friends who are very close and frequent the same group of friends, they probably think that their news feed is similar but this is not the case at all. They live in two different worlds where computers choose what to show them according to what matches you the best. Roger McNamee says that it is as if there were 2 billion different news feeds. Everyone lives in his own reality with his own facts. This is an interesting sentence because it allows us to introduce the next point.

B.11.14. Polarization of society and loss of objectivity

In continuity with the last sentence of the previous subsection, Roger McNamee explains that with the time we spend on these networks, we end up believing that everyone agrees with us because our news feed only shows that. When you're in that state of mind, it's very easy to manipulate you the way a magician would : a magician shows you cards and says pick one but what you don't realize is that he has it all planned. We end up picking the card he wants.

Facebook works the same way : they may tell you *chose your friends, chose the articles you want* but it's an illusion. Just like a magician, it's Facebook that elaborates everything from A to Z.

Rashida Richardson joins Roger McNamee in explaining that each person is informed with different facts and that when this phenomenon occurs on a large scale, it becomes impossible to take into account or even read information that does not agree with our world view. This means that we are no longer objective and constructive individuals.

If we look at politics, and we start looking at the other political clan (opposed to our ideas), we wonder how they can be so stupid? We say to ourselves *look at all this information I see every day, how come they don't see the same thing*. The answer is simple, they do not receive the same information [Justin Rosenstein]. To give you some graphs, in chapter 1.2, you will find the the evolution of the political polarization in the U.S that comes from the Netflix documentary (the source of those graphs are : Pew Research Center). On those graphs, it's interesting to see that between 1994 and 2004, the polarization was small and has remained almost unchanged. Then between 2004 and 2007, social networks arrived and we see small change. After 2007, we see that the polarisation increases more and more.

The next subsection will talk about the recommendations that these algorithms make to the users. This is directly related to this subsection on the polarization of society.

B.11.15. The problem of recommendations

Guillaume Chaslot is worried that the recommendation algorithm he contributed to is polarizing society. He explains that creating an algorithm that causes this polarization is very effective in keeping people connected. We often think that the goal of the algorithm is to give us what we want, but it's not true. At YouTube, its purpose is to make us dive as deep as possible in YouTube and the goal is to find out which video category best matches our interests. If you start watching one of these videos, others of the same style will be recommended to you constantly. This is what we call *Rabbit Hole Effect* which is very powerful.

YouTube is powerful to manipulate people and make them believe completely crazy things. For example, Basketball player Kyrie Irving apologized after saying that the earth was flat after watching a series of videos on YouTube about it [Tristan Harris] [315]. Guillaume Chaslot

adds the fact that the recommendation of the video saying that the earth is flat has been recommended millions of times by the algorithm. The problem is that the algorithm is becoming more and more intelligent and today, it manages to convince people that the earth is flat.

This leads us to the issue of fake news and the propagation of crazy theories. We will discuss this in the next subsection.

B.11.16. Fake news

An example of fake news spreading a completely crazy theory is the one that happened in December 2016 : there was the #pizzagate. Renee DiResta explained that the theory was that ordering a pizza was code for ordering a human being and that there was human trafficking. As these groups grew on the platforms, these platforms (Facebook, ...) recommended these groups more and more to their users. For example, if the social network sees that a person has shown interest in the anti-vaccine movement, this shows the Facebook algorithm that this is an individual who is sensitive to any form of conspiracy movement. The recommendation system would therefore lead them to these groups. The climax of the #pizzagate is when a person arrived armed in the pizzeria pizzeria Comet Ping Pong to liberate the human beings in the basement when there was not even a basement. This is a perfect example of a conspiracy theory that has been propagated all over social networks. The worst part is that it was their referral service that created this. It is also interesting to know that the owner of this restaurant, Comet Ping Pong, is known to be close to the Democratic Party and in particular to Tony Podesta, brother of Hillary Clinton's campaign manager [292].

If we focus on fake news in general, Tristan Harris explains that MIT did a study and showed that fake news spreads 6 times faster than real information on Twitter! There is a real concern about the balance of propagation between the information which is due to these recommendation algorithms. In this study, they also found that a fake news has 70% more chance to be retweeted than a true news but that's not all : it also takes true stories about six times as long to reach 1,500 people as it does for false stories to reach the same number of people. When it comes to Twitters cascades or unbroken retweet chains, falsehoods reach a cascade depth of 10 about 20 times faster than facts. And falsehoods are retweeted by unique users more broadly than true statements at every depth of the cascade [288].

We have therefore created a system that clearly favors the propagation of false information

because it generates more money for these companies than the real information [Sandy Parakilas]. It's a business model that works on misinformation. The more unregulated the messages, the higher the profiles. It becomes a platform where the highest bidder wins [Tristan Harris]. Between January 18 and 26, queries including the keywords beer corona virus increased by 2300% worldwide. People stopped buying Corona beer because of fake news saying that the coronavirus comes from this beer. The search beer virus has increased by 744%, while beer coronavirus has increased by 3233% [243][332]. To stay on the topic of covid, there was also a spread of the rumor that microchips were injected with the vaccines in order to track us[194] or that this virus was launched on purpose by the government to decrease the population...[80]

Tristan Harris explains that Russia and China have taken advantage of these rumors about covid to launch conspiracies. No one is able to separate the true from the false anymore. In China, Chinese agents helped spread messages that sowed virus panic in U.S., officials say. American officials were alarmed by fake text messages and social media posts that said President Trump was locking down the country [141]. What happened during the covid is only one example. Networks tend to fill up with all kinds of rumors and the end result is that we don't know what is true anymore, no matter what the subject is [Tristan Harris]. Roger McNamee explain that the problem with Facebook is that it is a very effective tool of persuasion, the best one ever. Imagine the effects it can have in the hands of a dictator or an authoritarian. If it takes us the desire to control the population of our country, there is no tool more effective than Facebook !

B.11.17. Politics

The fact that the government and other political parties use social networks to promote their ideas is very dangerous. There are repercussions in real life. One notable example is what happened in Myanmar: when Burmese people think about the Internet, they think mainly about Facebook and when they buy a new cell phone, the store installs Facebook directly and creates an account for them. So as soon as they have their cell phone, the first thing they know how to use is Facebook. Facebook has given the Burmese regime a way to manipulate public opinion, so they could spread hatred towards the Muslim population. There have been several crimes (murder, burning of villages, ...). This forced 700 000 Rohingya to flee the country [Cynthia M. Wong]. We see that Facebook can be used by third parties to easily manipulate

populations.

If you want to manipulate an election, you just have to find conspiracy groups on Facebook with hundreds of people who think that the earth is flat or that man has never walked on the moon. Then, we can ask Facebook to give us 1000 like them. They can bring us thousands of people in the same genre. On our side, we just have to bombard them with new theories. So it's pretty simple [Tristan Harris]. For Justin Rosenstein, algorithms and manipulative politicians are becoming experts in mental manipulation. They are now able to create false information that seems real. They confuse us so that we believe them. We have less and less control over who we are and what we believe. Imagine a world where no one believes in reality, in the government, ...that's where we're headed [Tristan Harris]. We can see that there is a rise of the extreme right and left in politics in all countries.

Unfortunately, the countries most affected by this polarization are the countries that organize democratic elections and the most developed. In these (democratic) countries, all you need is money and you can then destabilize what you want by paying social networks [Tristan Harris]. Roger McNamee talks about the 2016 elections in the USA and explains the rigged elections with Trump is not hacking by the Russians! They just had to understand how the algorithms worked. They used tools that Facebook had created for their normal advertisers by diverting them to nefarious purposes. The problem with these elections is that the goal was not to know for whom one was going to vote but for which clan one supported [Tristan Harris]. Marco Rubio said at a rally: we are a nation where people don't talk to each other anymore, where people are no longer friends just because of their vote, where people have isolated themselves and only listen to the information that makes them right. Therefore, a country can therefore invade another without even having to invade its borders [Justin Rosenstein].

B.11.18. Distinguishing truth from falsehood

We are made to believe that the engineers at the origin of this network will be able to solve the problem of fake news. That's a lie. AIs are not able to tell if a piece of information is false or not, they don't know how to distinguish them. They don't have the technology for that [Cathy O'Neil]. Moreover, if we can't agree on what the truth is and on the fact that the truth can exist, we're screwed. Because if we don't agree on what the truth is, we will never be able to solve the rest of the problems [Tristan Harris].

B.11.19. Are they a threat ?

In the end, it is not technology that threatens our existence but the fact that it is able to bring out the worst in human society. It is that dark side in all of us that is the real threat. It can create complete chaos, isolation, uncertainty, polarization, populism, manipulation, more distractions to keep us away from the real issues, ... but that's what today's society is all about! It affects everyone, even the people who don't use these platforms [Tristan Harris]. Tim Kendall explains that what scares him is that a civil war breaks out. For Jaron Lanier if we don't change anything, within 20 years, we will have destroyed our civilization and missed our objectives to save the planet. We will degrade our democracies which will become autocratic regimes. The world economy will also be destroyed and we will not survive.

Will we be the last generation to know what it was like to live before this illusion was put in place ? How do we get out of the matrix when we don't know it exists ? [Tristan Harris]

Tristan Harris makes a nuance in his words by saying that everything is not always black or white, we should not say that technology is busy destroying us. The disturbing thing is that we are living at the same time a utopia and a dystopia because it is fabulous to be able to press a button and 30 seconds later have a car that takes us where we want. Justin Rosenstein created the like button on Facebook with a team. In creating this button, he explains that their goal was to spread love and positivity in the world. But it did not take the turn they wanted (depression of some people for lack of likes, ...). Basically, none of this was bad, the real problem is the business model [Joe Toscano]. The problem is that their revenue has to increase and so they don't want to review their business model and change their algorithm [Alex Roetter]. They are therefore trapped by their own model because it is complicated to get out of it due to the fact that they want to have more and more money [Tristan Harris]. We live in a world where a tree and a whale have a greater financial value dead than alive... As long as our economy works this way and companies are not regulated, they will continue to kill trees and whales even if we know that this destroys the earth. It is a short term vision where we run after money [Tristan Harris].

According to Jaron Lanier, Joe Toscano and Sandy Parakilas, these companies are in a monopoly and are controlled by finance (always wanting more money). That's why regulations are needed but for Roger McNamee laws are lagging behind.

B.11.20. Some interesting ideas

Joe Toscano proposes as a solution to put in place a tax on the amount of data they collect.

For Justin Rosenstein, technology does not work with the laws of physics, nothing is set in stone. It is human beings who have decided what to do and these human beings can change this technology. The question is to know if we are able to admit that the negative repercussions of these networks are directly linked to our way of conceiving these networks. It is our right to ask that these products are made in an ethical and human way [Tristan Harris].

In history, every time there has been something positive, it is thanks to people who have questioned our way of doing things, who have criticized [Jaron Lanier].

C. Explanations

C.1. Fake news

C.1.1. What's a fake news ?

Depending on our point of view, different definitions can exist but we will see that putting a precise definition is not an easy thing. First of all, it is necessary to distinguish misinformation and disinformation.

1. **Misinformation** : is false information that is spread, regardless of intent to mislead [201].
2. **Disinformation** : refers to false information that is intentionally intended to mislead by distorting the facts [52]. So disinformation is misinformation that is intentionally spread. This includes propaganda, manipulation, ... [201]

Although the distinction between misinformation and disinformation is precise and clear, everyone agrees that the notion of fake news is subject to various interpretations [349]. Some are very precise in their criteria while others decide to include a large range of different type of false information as rumors, meme, ... For example, politicians may label information that does not support their position as a fake news and conversely, label information that does support their position as true [346]. It is obvious that this is not a good way to define what is fake news. That's why some people prefer to use the term false and true information instead of fake news [349] [68].

By analyzing several articles we realize that there are some subtle differences in the definition of what is fake news. However, if we summarize the definition, what stands out is that fake news refers to fabricated information that mimics news media content in its form and verifiable

false [238] [349]. And so some exclude from their definition rumors, conspiracy theories, satire, false statements by politicians, reports that are slanted or misleading but not outright false [68].

The way to define what is fake news or what is included in the notion of false information has, of course, an impact on the result of the different studies. Indeed, if the methodologies used by a study is to label their data as true or false. A data can be labeled differently depending on the definition used.

C.1.2. Rumor cascades

Figure C.1.: Complementary cumulative distribution functions (CCDFs) of false political and other types of rumor cascades. [346]

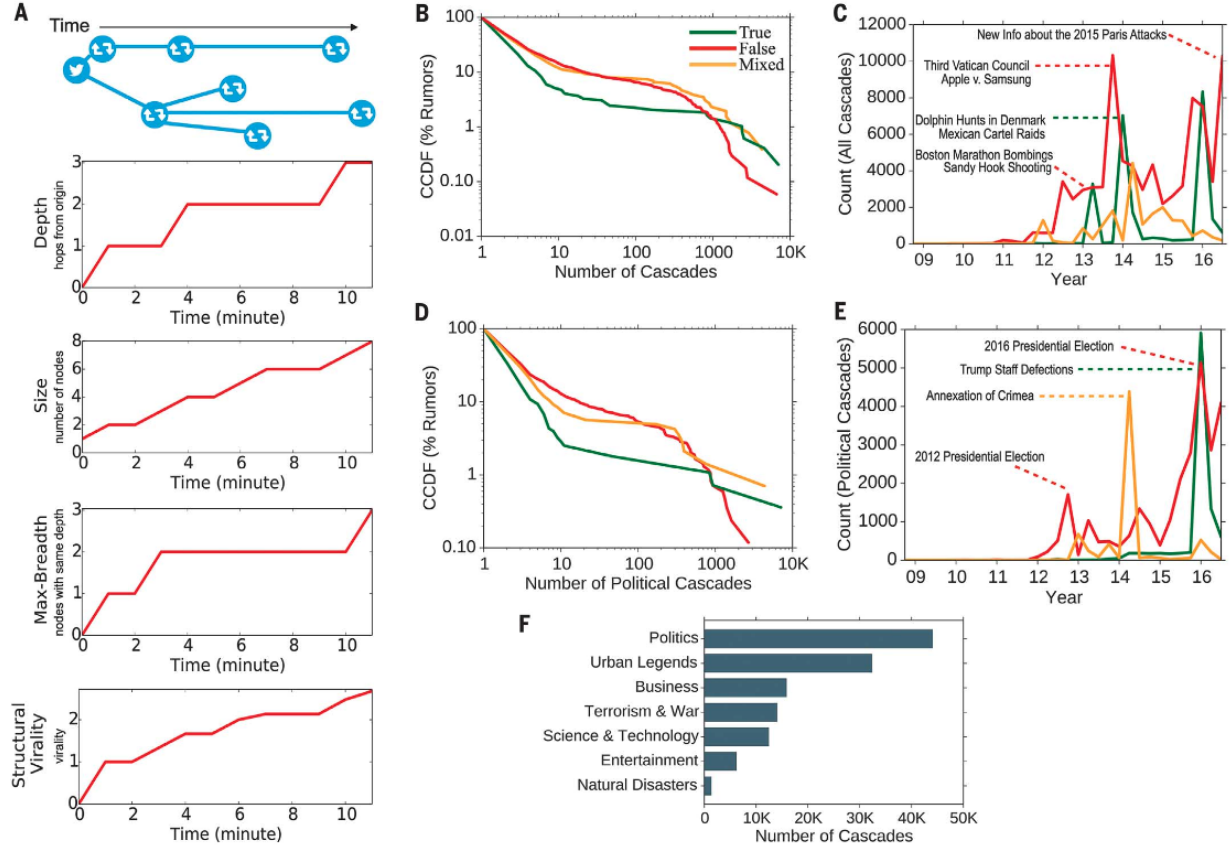


Fig. 1. Rumor cascades. (A) An example rumor cascade collected by our method as well as its depth, size, maximum breadth, and structural virality over time. "Nodes" are users. (B) The complementary cumulative distribution functions (CCDFs) of true, false, and mixed (partially true and partially false) cascades, measuring the fraction of rumors that exhibit a given number of cascades. (C) Quarterly counts of all true, false, and mixed rumor cascades

that diffused on Twitter between 2006 and 2017, annotated with example rumors in each category. (D) The CCDFs of true, false, and mixed political cascades. (E) Quarterly counts of all true, false, and mixed political rumor cascades that diffused on Twitter between 2006 and 2017, annotated with example rumors in each category. (F) A histogram of the total number of rumor cascades in our data across the seven most frequent topical categories.

Figure C.2.: Complementary cumulative distribution functions (CCDFs) of true and false rumor cascades. [346]

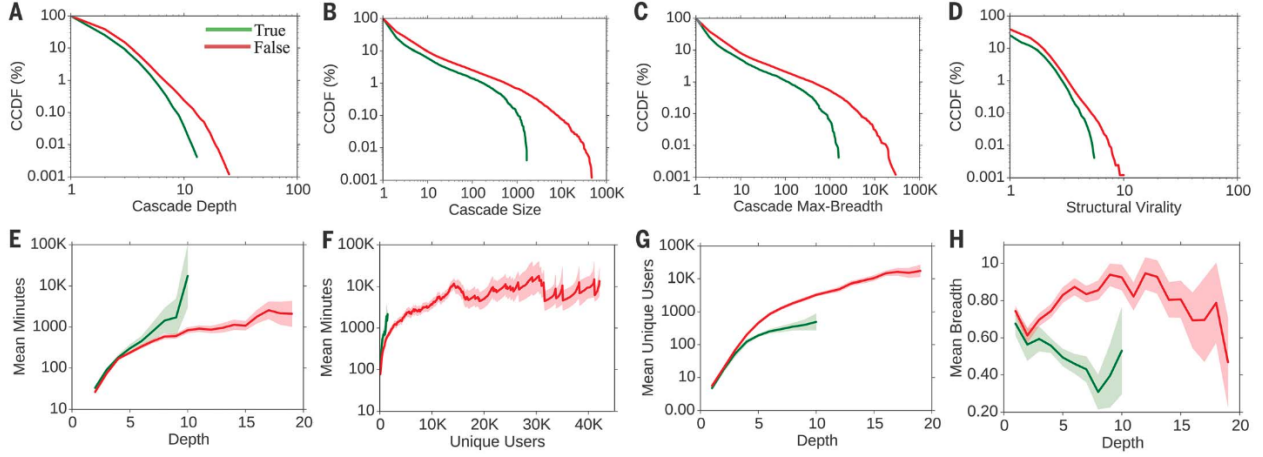


Fig. 2. Complementary cumulative distribution functions (CCDFs) of true and false rumor cascades. (A) Depth. (B) Size. (C) Maximum breadth. (D) Structural virality. (E and F) The number of minutes it takes for true and false rumor cascades to reach any (E) depth and (F) number of unique Twitter users. (G) The number of unique Twitter

users reached at every depth and (H) the mean breadth of true and false rumor cascades at every depth. In (H), plot is lognormal. Standard errors were clustered at the rumor level (i.e., cascades belonging to the same rumor were clustered together; see supplementary materials for additional details).

Figure C.3.: Rumor cascades. [346]

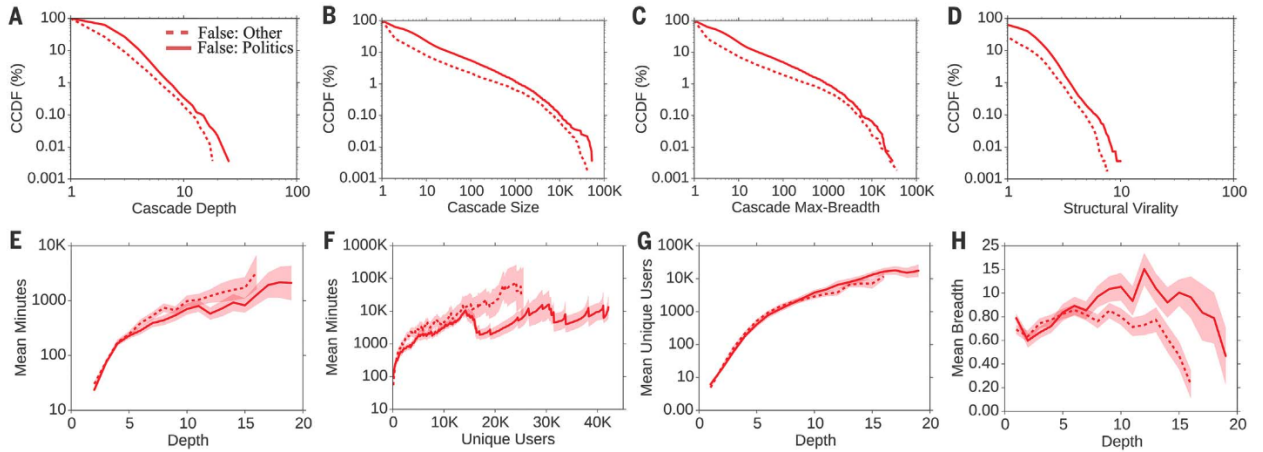


Fig. 3. Complementary cumulative distribution functions (CCDFs) of false political and other types of rumor cascades. (A) Depth. (B) Size. (C) Maximum breadth. (D) Structural virality. (E and F) The number of minutes it takes for false political and other false news cascades to reach

any (E) depth and (F) number of unique Twitter users. (G) The number of unique Twitter users reached at every depth and (H) the mean breadth of these false rumor cascades at every depth. In (H), plot is lognormal. Standard errors were clustered at the rumor level.

Figure C.4.: What Political News and Information Were US Twitter Users Sharing during the Election? [186]

<i>Type of Source</i>	<i>N</i>	<i>% of Subset</i>	<i>% of Total</i>
Professional News Content			
Major News Brands	196,697	77	
Minor News Brands	60,055	23	
Subtotal	256,752	100	20
Professional Political Content			
Political Party or Candidate	121,323	93	
Experts	6,350	5	
Government	2,294	2	
Subtotal	129,967	100	10
Polarizing and Conspiracy Content			
Junk News	203,591	79	
WikiLeaks	48,068	19	
Russia	7,683	3	
Subtotal	259,342	100	20
Other Political News and Information			
Other Political	97,900	59	
Citizen, Civic or Civil Society	34,935	21	
Political Humor or Entertainment	30,021	18	
Religion	2,124	1	
Political Merchandise	2,067	1	
Subtotal	167,047	100	13
Other			
Social Media Platform	195,470	42	
Not Available	241,688	52	
Other Non-Political	25,164	5	
Subtotal	462,322	100	36
TOTAL	1,269,736		100

*Source: Authors' calculations from data sampled 1-11/11/16. For the full list of hashtags used to capture this sample see Memo 2017.1).*¹¹

C.2. Discover your emotions

With TikTok and Instagram, it's easy to discover what emotion you have. With the time you spend on each video, they can adapt what will be the next ones. Progressively, they will

discover how you feel, what makes you feel that way and what interests you at the moment. The aim with this is of course to keep you on the platform by playing with the emotions they have discovered.

For example on Instagram, if my goal is to find out if you are currently sad, this is an information that I can not get directly but indirectly by showing you sad reels¹ and see how you react to it: Do you spend more time on these reels? Do you scroll through more posts? ... The second objective is to know the subject that generates this sadness. Again, I just need to put different sad reels on different subjects and analyze your behavior. Progressively and imperceptibly, I infiltrate your private life while you are scrolling, innocently, on the platform. Then, when I have captured the subject of your emotion, I just have to gradually scrape more and more information about what you feel or simply amplify/diminish your emotions with what I propose to you. You have simply become a puppet where I pull the strings.

C.3. Dopamine vs serotonin

Dopamine and serotonin are inversely correlated in our body. If one increases, the other decreases and vice versa [163]. Dopamine is a hormone that gives us instant pleasure and satisfaction (excitement and short term) while serotonin is a hormone that gives us long term happiness (peace and long term) [13][163]. The activities that promote the secretion of serotonin are: having a healthy lifestyle, eating healthy, contributing to the improvement of society, ... The fact that certain substances are considered as drugs or that we have addictions is partly due to dopamine. Any substance, activity, action that boosts dopamine makes us likely to become dependent on it (which is not the case for serotonin) [163][5]. An overdose of dopamine or a bad balance dopamine/serotonin can lead to several consequences:

- A long term dopamine overload can lead to a desensitization of our receptors, which consequently, makes us want more and more dopamine and especially, when we return to a normal level, we feel bad, anxious, As said before, this hormone plays on our concentration, our memory, our sleep, ... and therefore we enter a vicious circle where

¹Reels refer to brief video clips, usually lasting less than one minute. Unlike the traditional scrolling experience found on platforms such as Facebook, where users navigate through different types of content, such as text, images, and articles, scrolling through reels entails an endless stream of concise videos. This format provides users with a continuous flow of captivating video content [333][6][97].

we can not do our work, sleep well, ... [13].

- If dopamine increases, serotonin decreases. So, when we have over-solicited our dopamine for too long, we have a loss of happiness because we have less serotonin.
- One of the main consequences of a dopamine overdose is that it can lead to schizophrenia and an increase of aggressiveness [13][77].

D. Survey

D.1. Surveys - UCLouvain student

This survey was exclusively intended for students of the UCLouvain and was anonymous. It is composed of 3 parts:

1. Basic personal information (4 questions).
2. Your use of networks (6 questions).
3. What are your sources of information on current affairs (15 questions)

The results are not representative of reality due to the number of responses and the fact that the majority of responses came from the EPL faculty and FIAL faculty (thus there is little diversity). However, it allows us to give some clues and to have an overview of certain questions that we could ask ourselves. Combine with other results from other studies. It becomes interesting.

Enquête - Réseaux sociaux

181 réponses

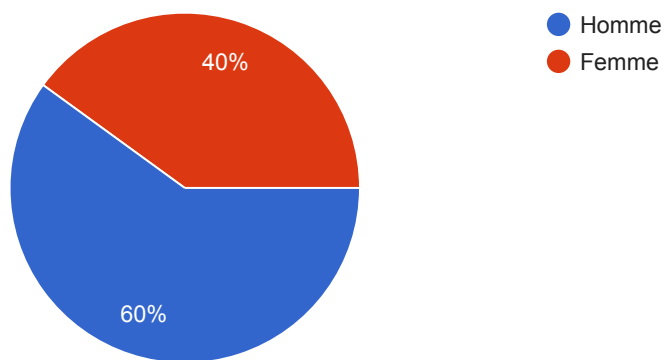
[Publier les données analytiques](#)

Informations personnelles

(1/4) Sexe :

180 réponses

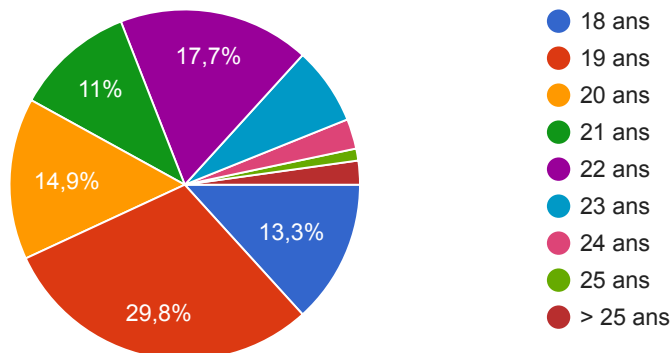
 Copier



(2/4) Age :

181 réponses

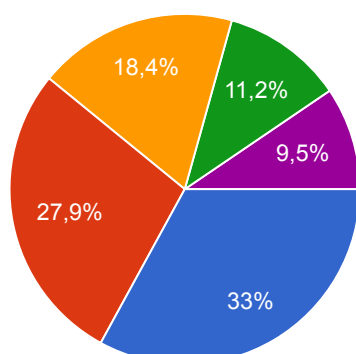
 Copier



(3/4) Année d'étude :

 Copier

179 réponses

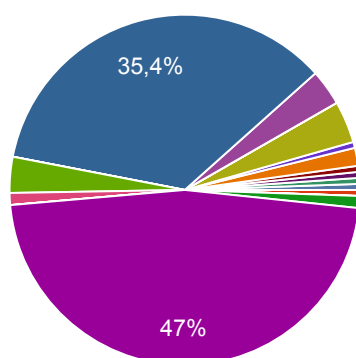


- Bachelier 1
- Bachelier 2
- Bachelier 3
- Master 1
- Master 2

(4/4) Faculté d'étude :

 Copier

181 réponses



- MEDE (Faculté de médecine...)
- FASB (Faculté de pharmacie...)
- FSP (Faculté de santé publiq...)
- FSM (Faculté des sciences d...)
- EPL (Ecole polytechnique de...)
- LOCI (Faculté d'architecture,...)
- AGRO (Faculté des bioingéni...)
- SC (Faculté des sciences)

▲ 1/3 ▼

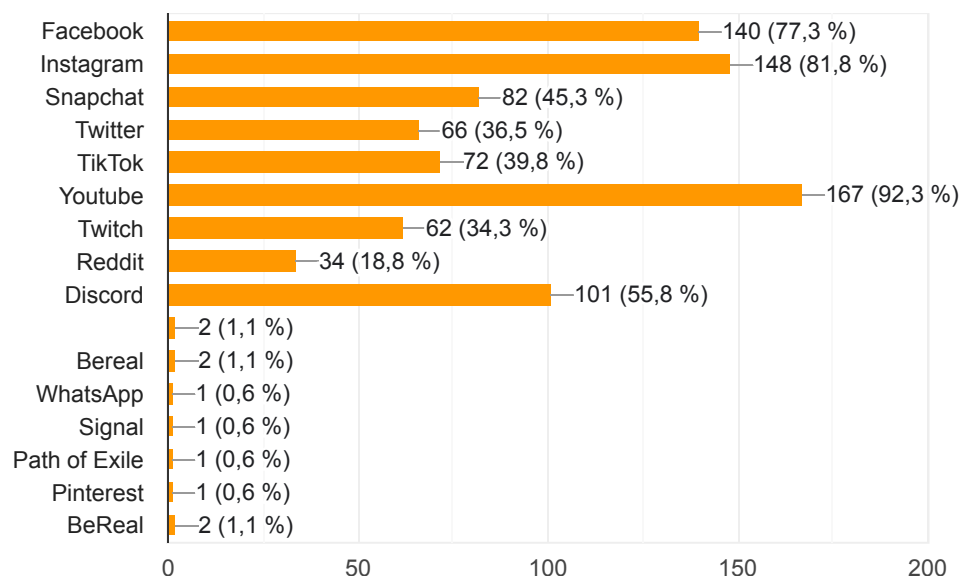
Votre utilisation des réseaux sociaux



(1/6) Quels réseaux sociaux utilisez-vous ?

 Copier

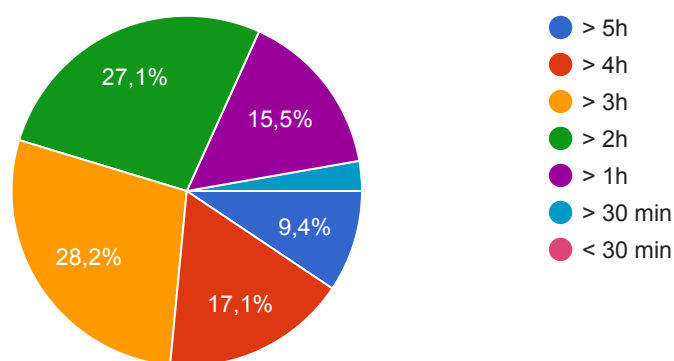
181 réponses



(2/6) Combien de temps passez-vous sur les réseaux sociaux par jour (estimation) ?

 Copier

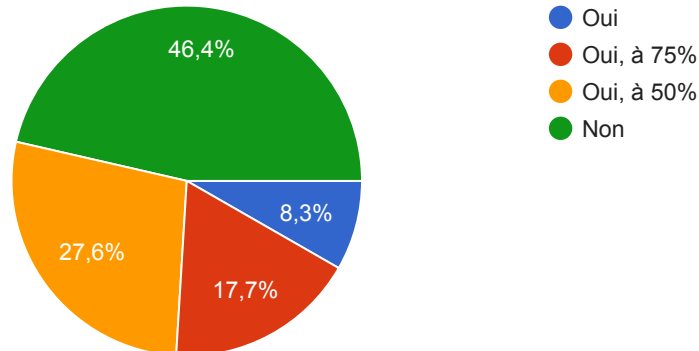
181 réponses



(3/6) Trouvez-vous que les réseaux sociaux indiquent clairement quelles informations sont récoltées à votre propos ?

 Copier

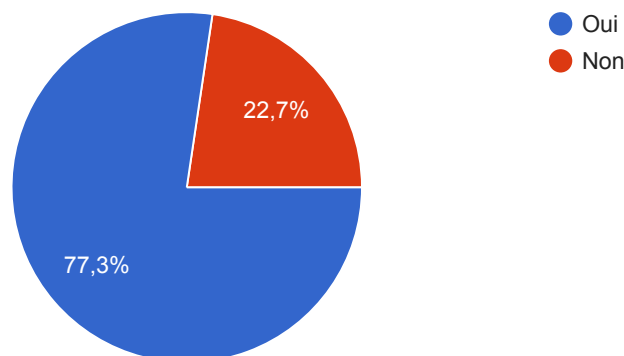
181 réponses



(4/6) Est-ce important (pour vous) de connaître les informations et le type d'informations qui sont récoltées à votre propos ?

 Copier

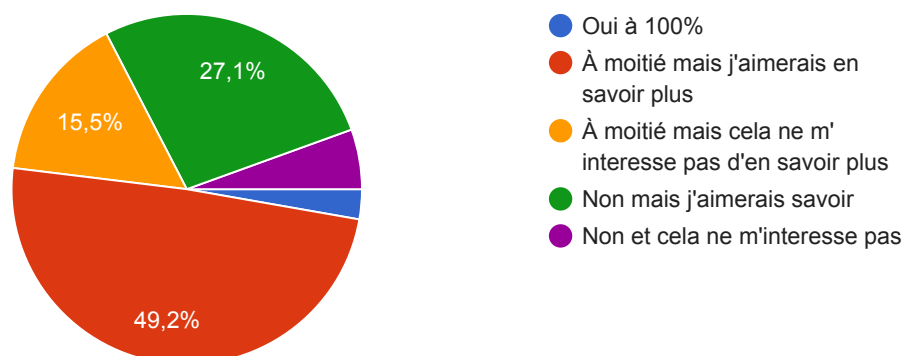
181 réponses



(5/6) Savez-vous ce qu'ils font avec vos données ?

 Copier

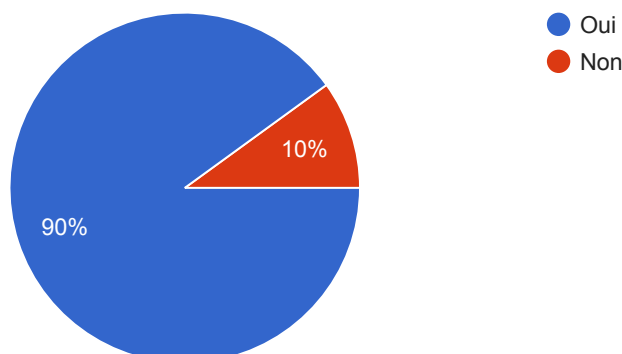
181 réponses



(6/6) Pensez-vous que cette récolte de données peut être utilisée à mauvais escient ?

 Copier

180 réponses



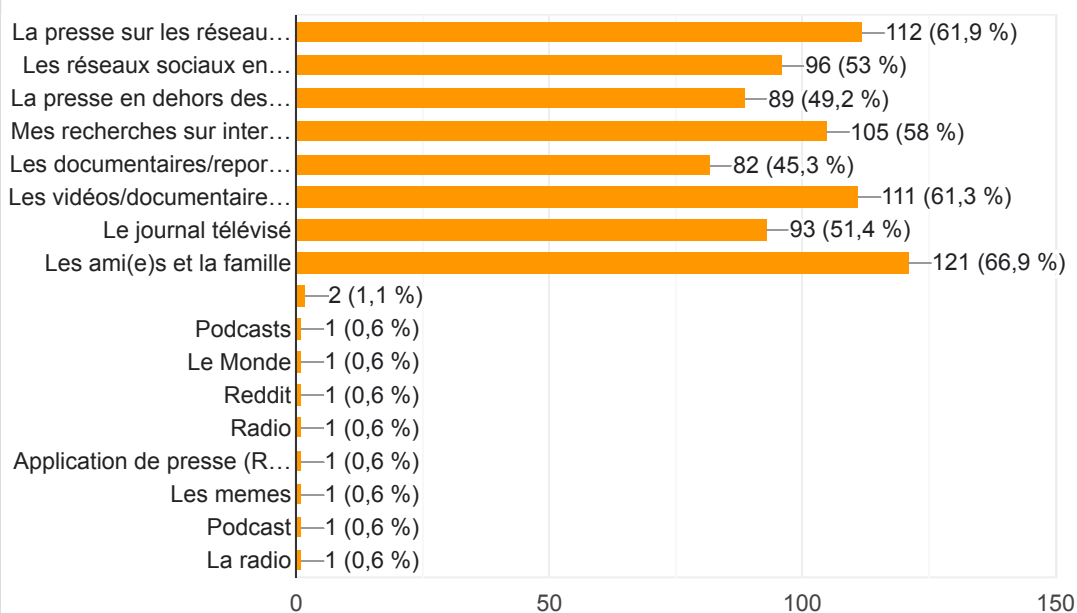
Quelles sont vos sources d'informations sur l'actualité ?

(1/15) Quelles sont vos sources d'informations sur l'actualité ?

 Copier

[Plusieurs réponses possibles]

181 réponses

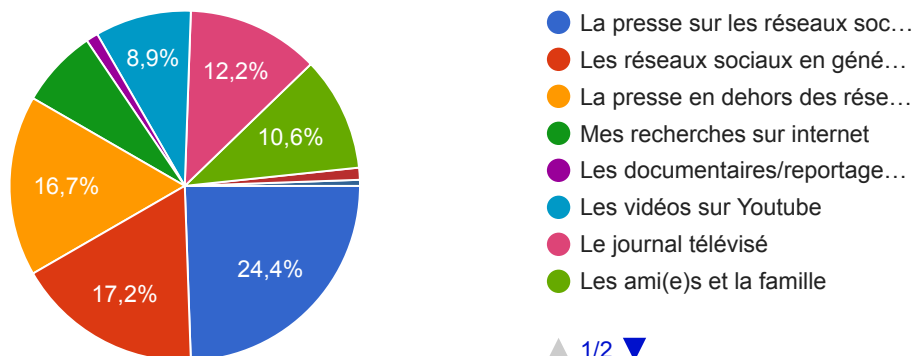


(2/15) Quelle est votre source principale d'information sur l'actualité ?

 Copier

[Une seule réponse]

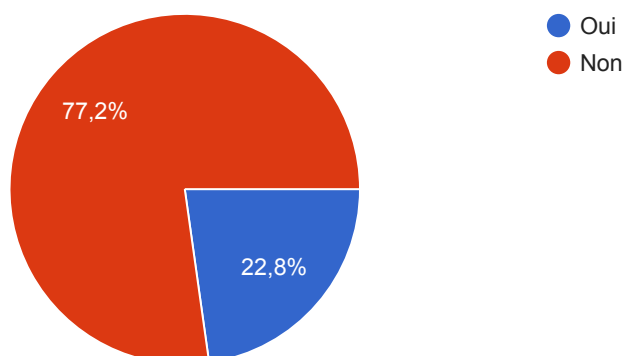
180 réponses



(3/15) Avez-vous un abonnement payant à une presse sur l'actualité (ex : Le Monde, Le Figaro, Le Soir, Charlie Hebdo, La libre, Libérations, RTBF, ...) ?

 Copier

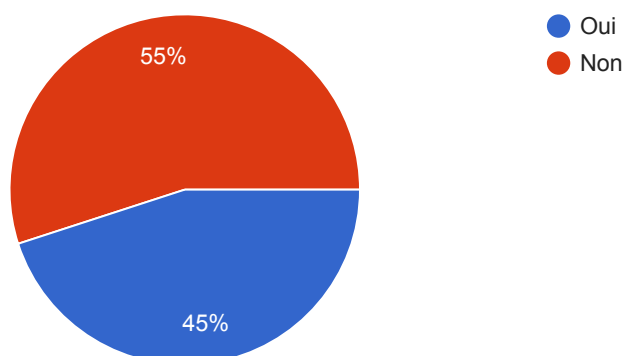
180 réponses



(4/15) Suivez-vous la presse (à propos de l'actualité) sur vos réseaux sociaux ? (ex : Le Monde, Le Figaro, Le Soir, Charlie Hebdo, La libre, Libérations, RTBF, ...)

 Copier

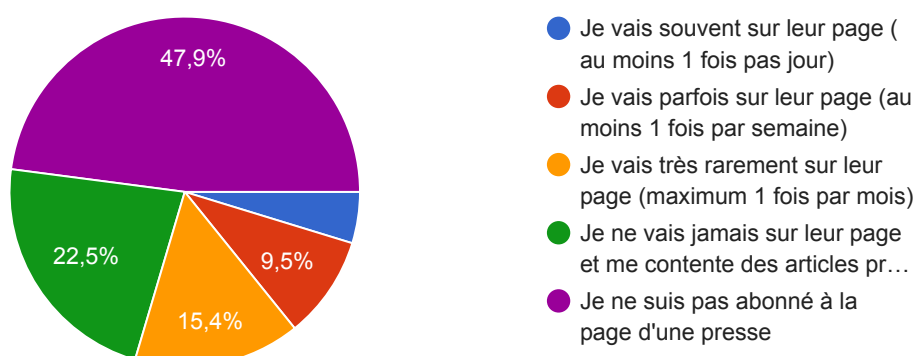
180 réponses



(5/15) Si vous êtes abonné(e) à la page d'une presse sur vos réseaux sociaux : allez-vous sur leur page ou lisez vous uniquement les articles proposés sur votre fil d'actualité ?

 Copier

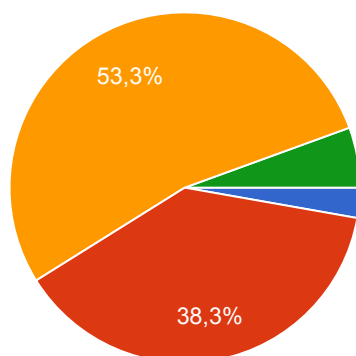
169 réponses



(6/15) Sur vos réseaux sociaux, si un article vous intéresse : lisez-vous uniquement les titres et les commentaires ou allez-vous voir l'article en cliquant sur le lien ?

 Copier

180 réponses

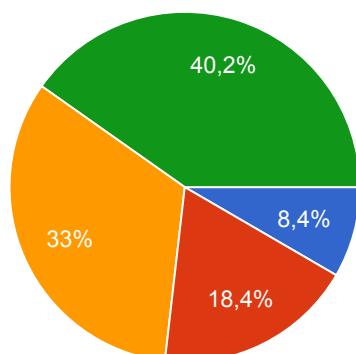


- Si l'article m'intéresse, je regarde uniquement le titre et les commentaires
- Si l'article m'intéresse, je regarde principalement le titre et les commentaires mais par...
- Si l'article m'intéresse, je vais toujours voir l'article en cliquant sur le lien
- Je n'est jamais d'article sur mes réseaux sociaux

(7/15) Regardez-vous les auteurs des articles que vous lisez (au minimum le nom) ?

 Copier

179 réponses



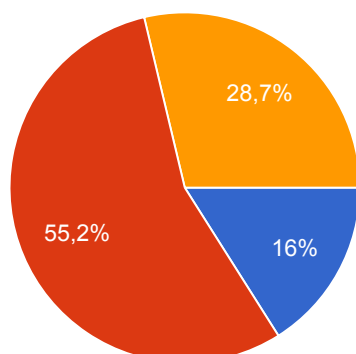
- Oui, tout le temps
- Parfois (1 fois sur 2)
- Rarement (moins d'une fois sur 2)
- Jamais



(8/15) Faites vous des recherches sur l'actualité ou lisez-vous majoritairement les articles qui vous sont proposés sur les réseaux sociaux ?

 Copier

181 réponses

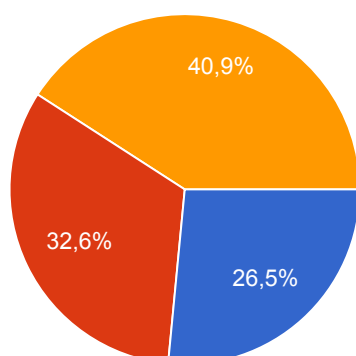


- Je fais majoritairement des recherches
- Je lis majoritairement ce que l'on me propose
- Je fais les deux à la même fréquence

(9/15) Lisez-vous plusieurs sources pour un même sujet d'actualité ?

 Copier

181 réponses

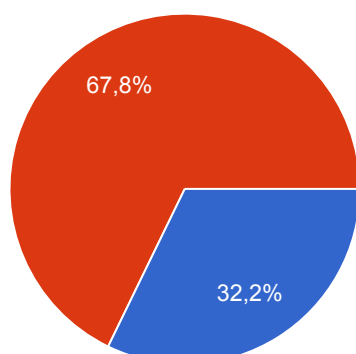


- Majoritairement oui.
- Majoritairement non.
- Oui et non à la même fréquence.

(10/15) Êtes-vous d'accord avec le fait qu'il faut payer certains articles pour avoir accès au contenu ?

 Copier

177 réponses



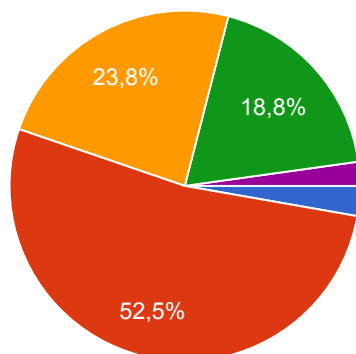
- Oui
- Non



(11/15) Pensez-vous partager les mêmes idées sur l'actualité, la politique, ... que votre entourage (ami(e)s, famille, ...) ?

 Copier

181 réponses

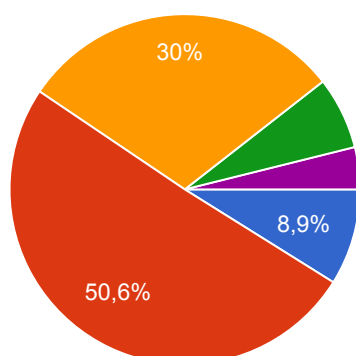


- Oui, à 100%
- Cela dépend du sujet mais majoritairement oui
- Oui, à 50%
- Cela dépend du sujet mais majoritairement non
- Non, jamais

(12/15) Faites-vous confiance aux journaux de masse tels que Le Monde, Le Figaro, le New York Times, la RTBF, ... ?

 Copier

180 réponses

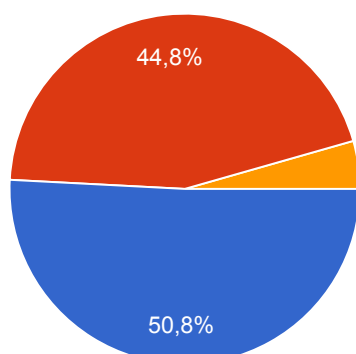


- Oui, à 100%
- Oui, à 75%
- Oui, à 50%
- Oui à moins de 50%
- Non

(13/15) Vous posez-vous la question, lorsque vous lisez un article, si c'est une "vraie information" ou une "fake news" ?

 Copier

181 réponses



- Oui
- Oui parfois (1 fois sur 2)
- Non

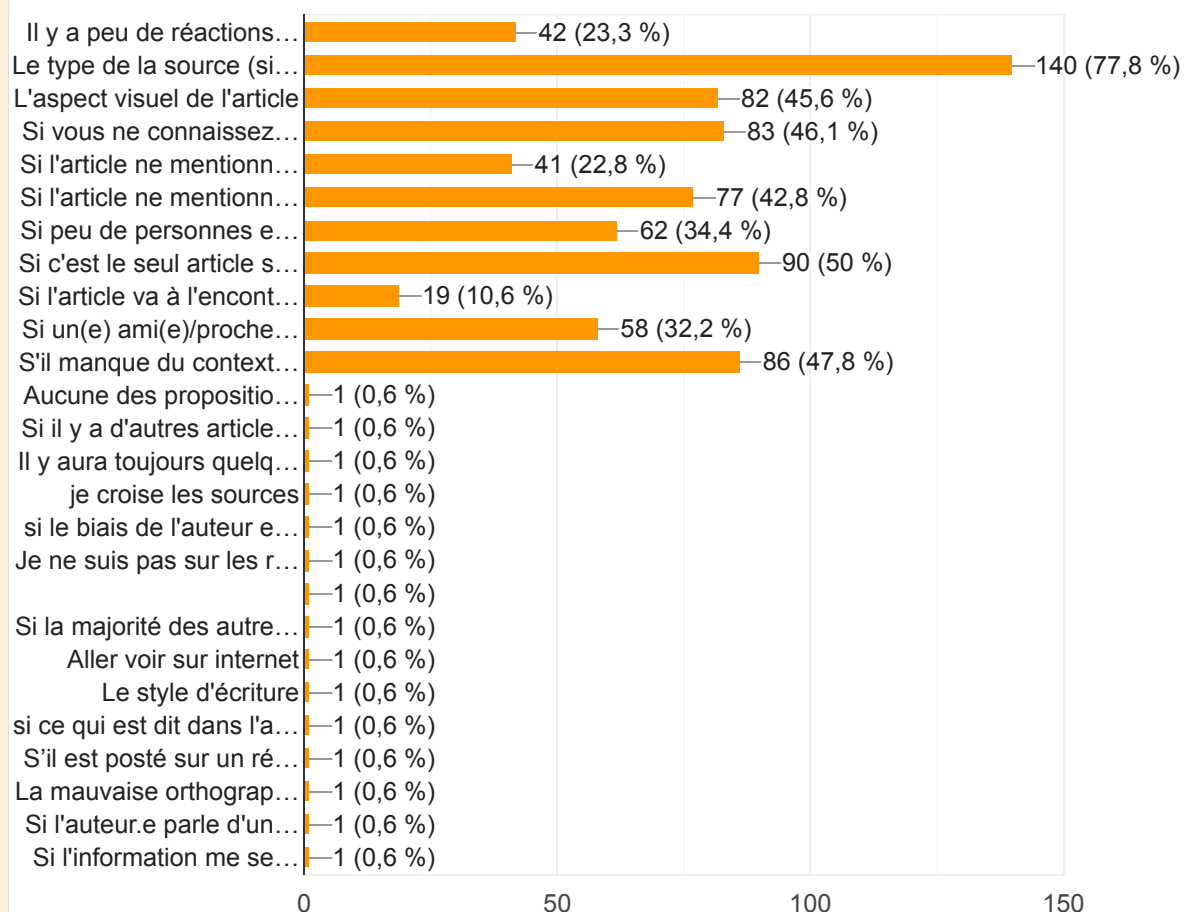


(14/15) Sur les réseaux sociaux, sur base de quels critères vous basez-vous afin de déterminer si un article est une "fake news" (autrement dit, une "fausse information") ? Attention, sélectionnez uniquement ce que vous faites et non ce que vous pensez devoir faire.

 Copier

[Plusieurs réponses possibles]

180 réponses

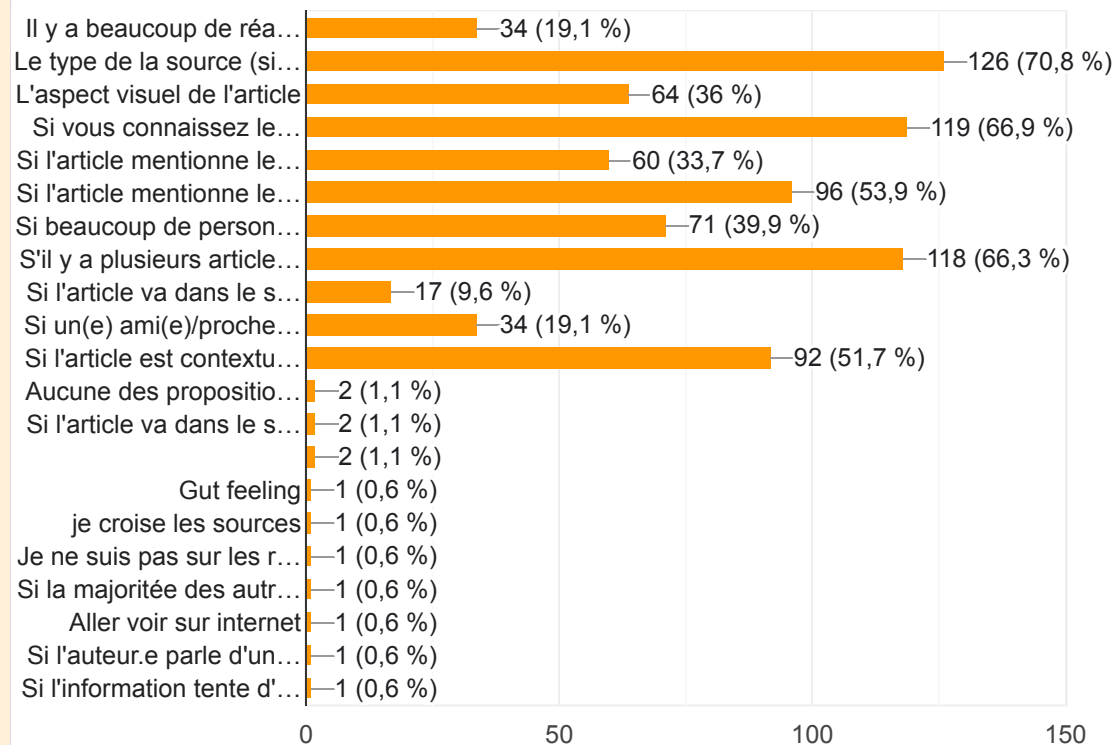


(15/15) Sur les réseaux sociaux, sur base de quels critères vous basez-vous afin de déterminer si un article est une "vraie information" ? Attention, sélectionnez uniquement ce que vous faites et non ce que vous pensez devoir faire.

 Copier

[Plusieurs réponses possibles]

178 réponses



Ce contenu n'est ni rédigé, ni cautionné par Google. [Signaler un cas d'utilisation abusive](#) - [Conditions d'utilisation](#) - [Règles de confidentialité](#)

Google Forms



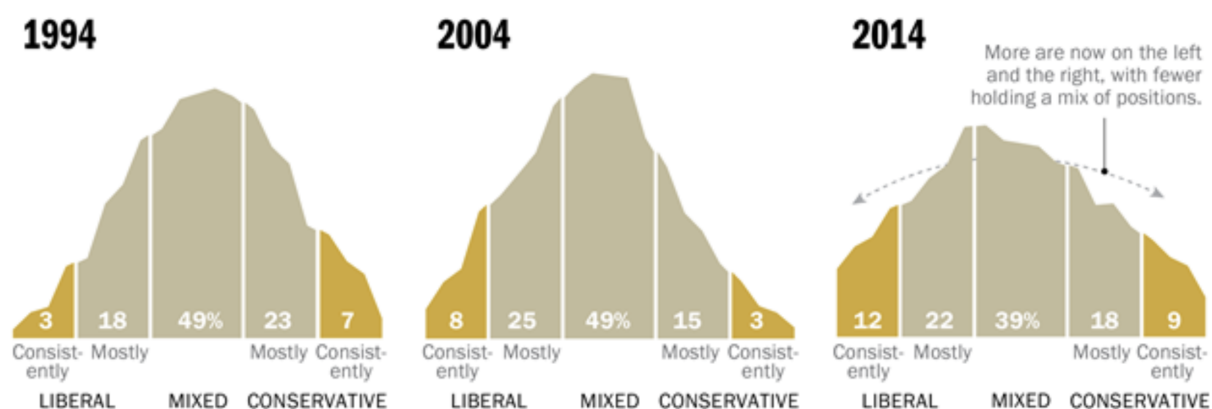
E. Collection of images, graphs, screenshots, etc.

E.1. Political polarisation in U.S

Figure E.1.: Growing minority holds consistent ideological views. [255]

Growing Minority Holds Consistent Ideological Views

On a 10-item scale of political values, % who are...

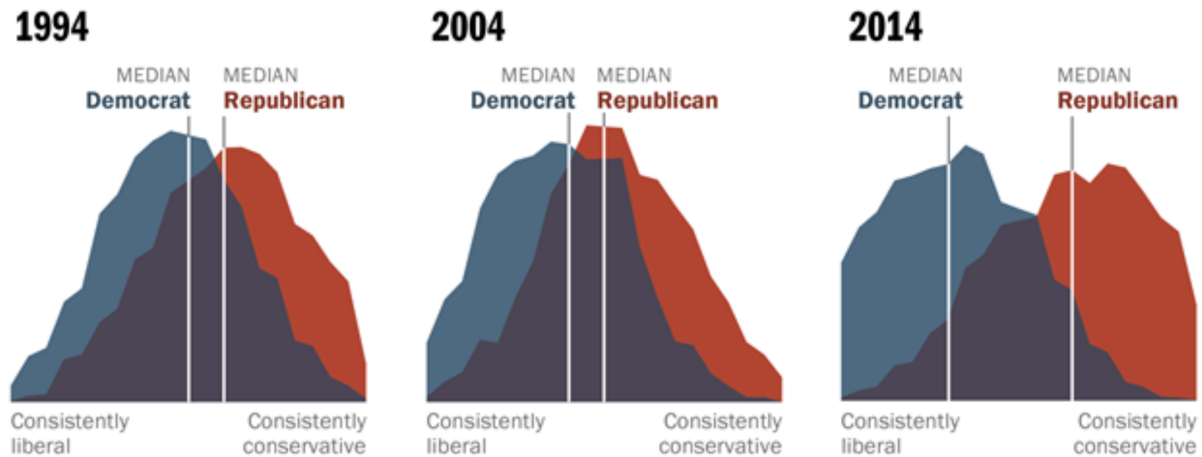


Source: 2014 Political Polarization in the American Public

Notes: Ideological consistency based on a scale of 10 political values questions. (See Appendix A for details on how the scale is constructed and how scores are grouped.)

PEW RESEARCH CENTER

Figure E.2.: Democrats and Republicans more ideologically divided than in the past. [255]

Democrats and Republicans More Ideologically Divided than in the Past*Distribution of Democrats and Republicans on a 10-item scale of political values*

Source: 2014 Political Polarization in the American Public

Notes: Ideological consistency based on a scale of 10 political values questions (see Appendix A). The blue area in this chart represents the ideological distribution of Democrats; the red area of Republicans. The overlap of these two distributions is shaded purple. Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents (see Appendix B).

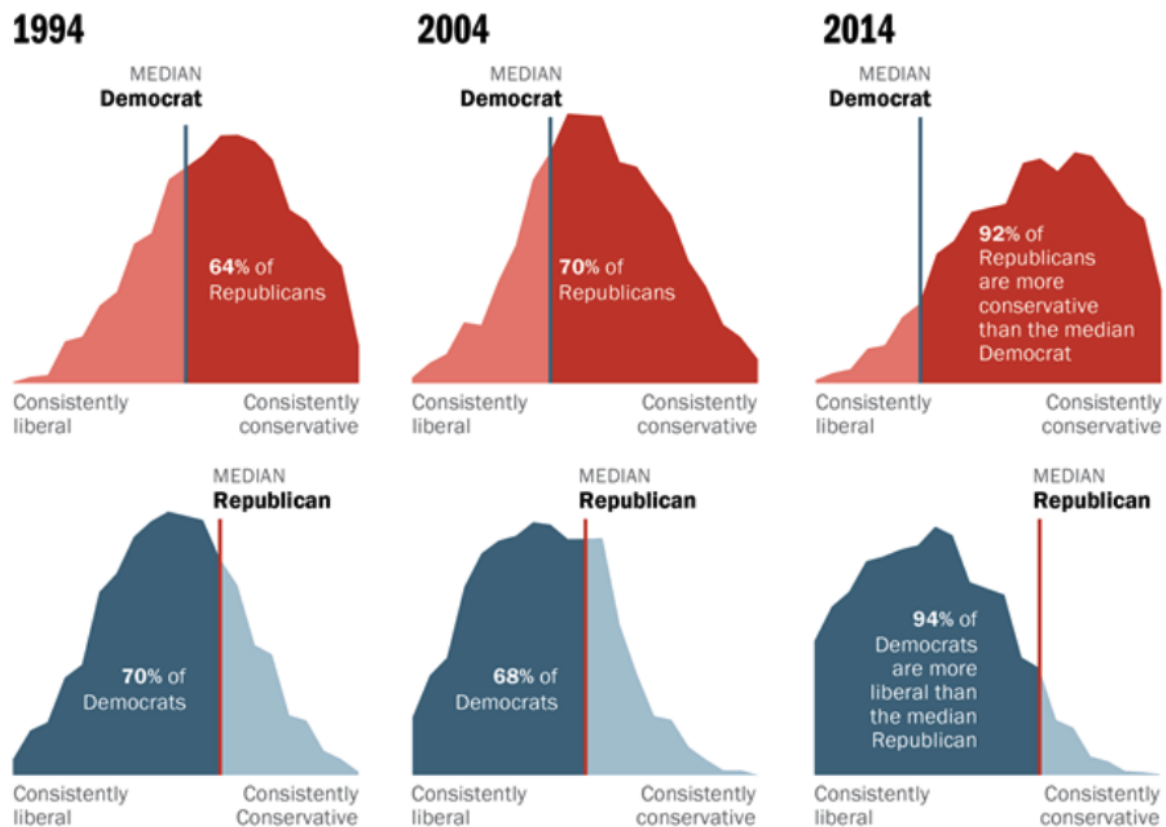
PEW RESEARCH CENTER

Figure E.3.: Republicans shift to the right, Democrats to the left. [255]

Republicans Shift to the Right, Democrats to the Left

Republicans Shift to the Right, Democrats to the Left

Distribution of Republicans and Democrats on a 10-item scale of political values



Source: 2014 Political Polarization in the American Public.

Notes: Ideological consistency based on a scale of 10 political values questions (see Appendix A). Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents (see Appendix B).

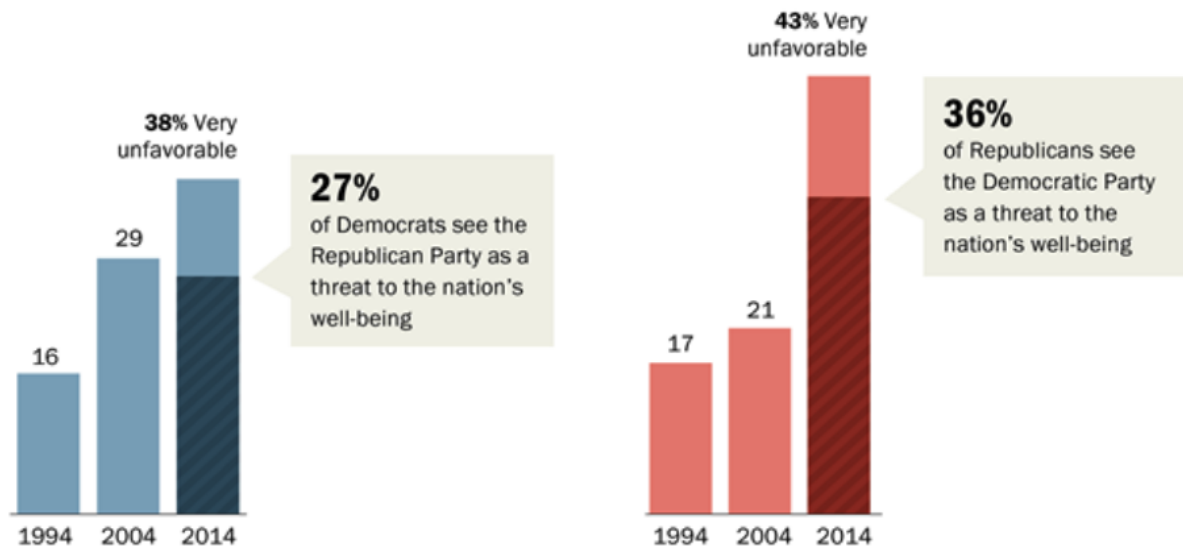
PEW RESEARCH CENTER

Figure E.4.: Beyond dislike: viewing the other party as a 'Threat to the nation's well-being.'
[255]

Beyond Dislike: Viewing the Other Party as a 'Threat to the Nation's Well-Being'

Democratic attitudes about the Republican Party

Republican attitudes about the Democratic Party

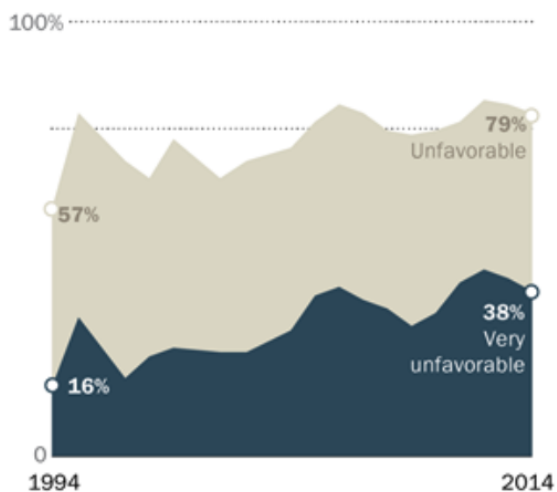
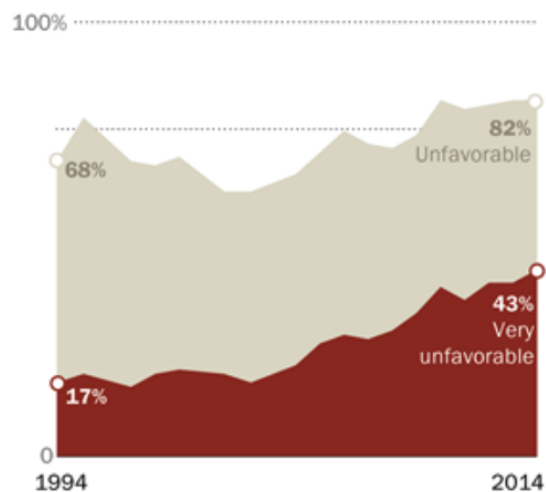


Source: 2014 Political Polarization in the American Public

Notes: Questions about whether the Republican and Democratic Parties are a threat to the nation's well being asked only in 2014. Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents (see Appendix B).

PEW RESEARCH CENTER

Figure E.5.: A rising tide of mutual antipathy. [255]

A Rising Tide of Mutual Antipathy**Democratic attitudes about the Republican Party****Republican attitudes about the Democratic Party**

Source: 2014 Political Polarization in the American Public

Note: Republicans include Republican-leaning independents; Democrats include Democratic-leaning independents.

PEW RESEARCH CENTER

E.2. Facebook advertising

E.2.1. Revenue generated by Facebook and Google through advertising

How much money does Facebook and Google generate through advertising?

Figure E.6.: Annual advertising revenue of Meta Platforms worldwide from 2009 to 2022 (in million U.S. dollars). [304]

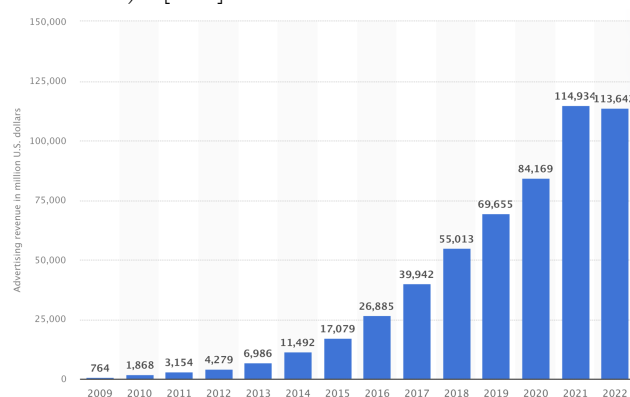


Figure E.7.: Advertising revenue of Google from 2001 to 2022 (in billion U.S. dollars). [335]

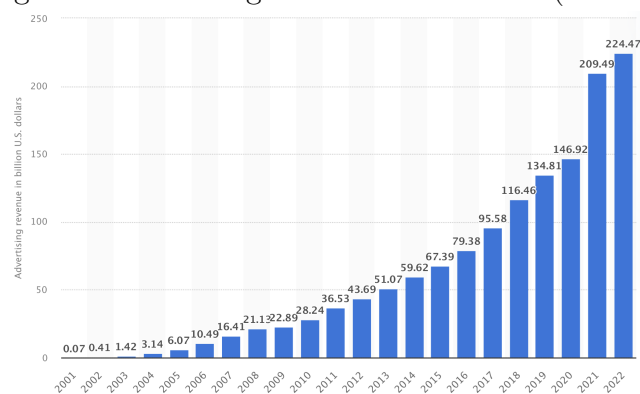
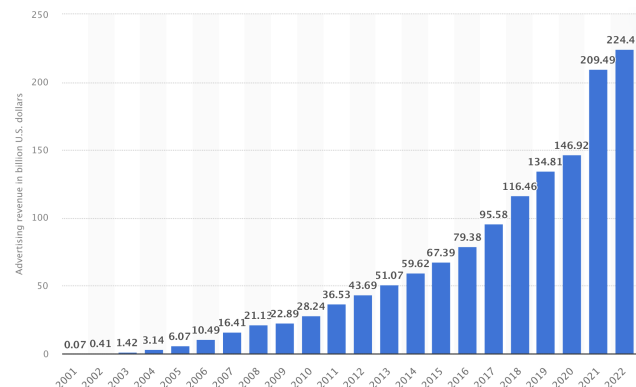


Figure E.8.: Advertising revenue and profit of Google from 2017 to 2022. [166]



E.2.2. Parameters selection for audience targeting in a Facebook advertisement

Figure E.9.: Possible objectives for advertising

Objectif

Quels r  sultats attendez-vous de cette publicit   ?

- ☒ **Automatique**
Laissez Facebook s  lectionner l'objectif le plus pertinent d'apr  s vos param  tres.
- ☐ **Recevoir plus d'appels**
Montrez votre publicit   aux personnes susceptibles d'appeler votre entreprise.
- ☐ **Encourager les visites en ligne**
Montrez votre publicit   aux personnes susceptibles de cliquer sur une URL qu'elle contient.
- ☐ **Recevoir plus de messages**
Diffusez votre publicit   aupr  s des personnes les plus susceptibles de vous envoyer un message.
- ☐ **Obtenez plus de mentions J'aime la Page**
Cr  ez une promotion pour aider plus de personnes    trouver et aimer votre Page.
- ☐ **Obtenir plus de prospects**
Utilisez un formulaire pour recueillir les coordonn  es de clients potentiels.

[Annuler](#) [Enregistrer](#)

Figure E.10.: Choice of the audience

Modifier l'audience

Sélectionnez le lieu, l'âge, le genre et les centres d'intérêt des [comptes de l'Espace Comptes](#) que vous voulez toucher avec votre publicité.

Genre ⓘ

Tous Hommes Femmes

Âge ⓘ

18 de 65 ans

Sélectionner une audience de moins de 18 ans limitera vos options de ciblage à cet âge et à certains lieux. [En savoir plus](#)

Lieux ⓘ

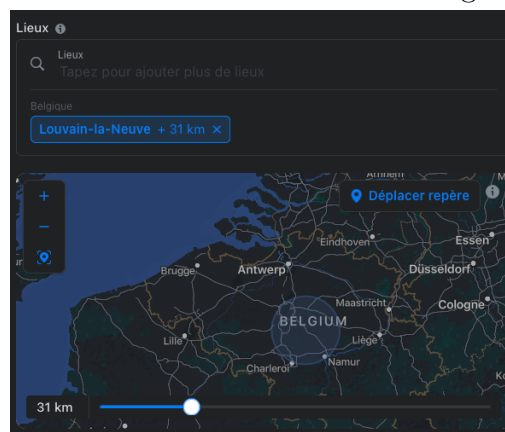
Lieux

Tapez pour ajouter plus de lieux

Belgique

Louvain-la-Neuve + 40 km x

Figure E.11.: Choice of the audience - geographic



The following 3 images do not contain an exhaustive list of possible choices.

Figure E.12.: Advanced targeting - demographic data

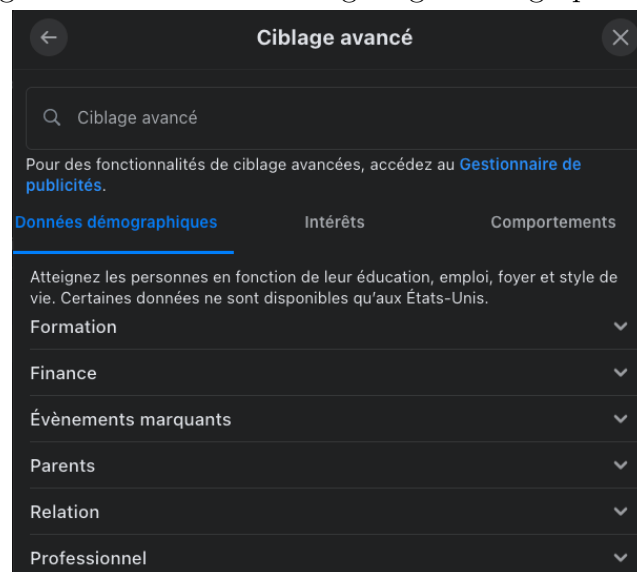


Figure E.13.: Advanced targeting - interests

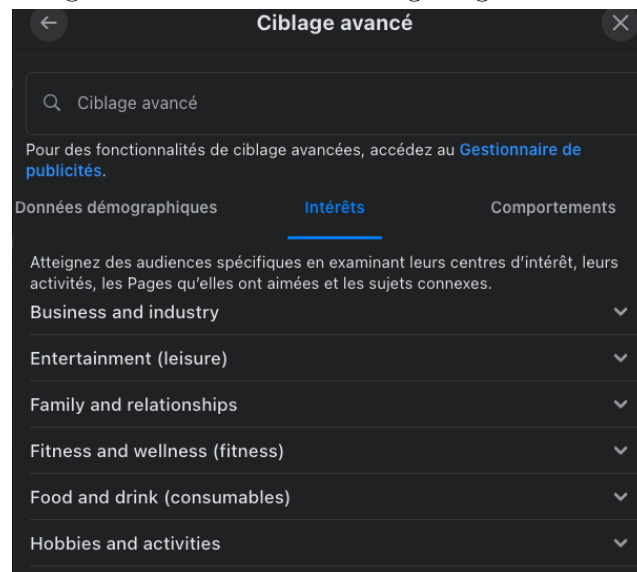
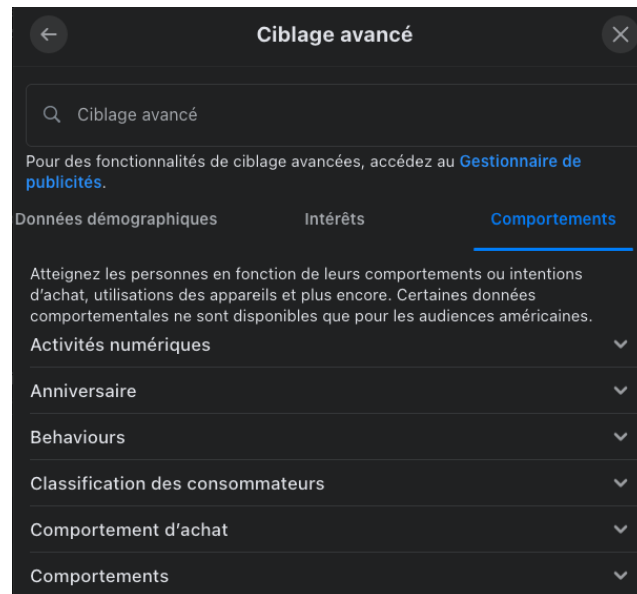


Figure E.14.: Advanced targeting - behaviors



E.3. Actions taken by social networks

E.3.1. Prevention messages from Youtube

In this section, we will show you different prevention messages about manipulation and recommendation that we have received from Youtube when we were browsing the platform privately with a computer. They appeared spontaneously because we did not know that Youtube had started to display this type of message. We do not know either how many people received these prevention. These messages appeared randomly between January 6 and March 15, 2023 with a concentration on January. Since then, we have not seen any more when we go on Youtube.

Figure E.15.: Date and time of the screenshot : Friday January 6, 2023 at 14:46

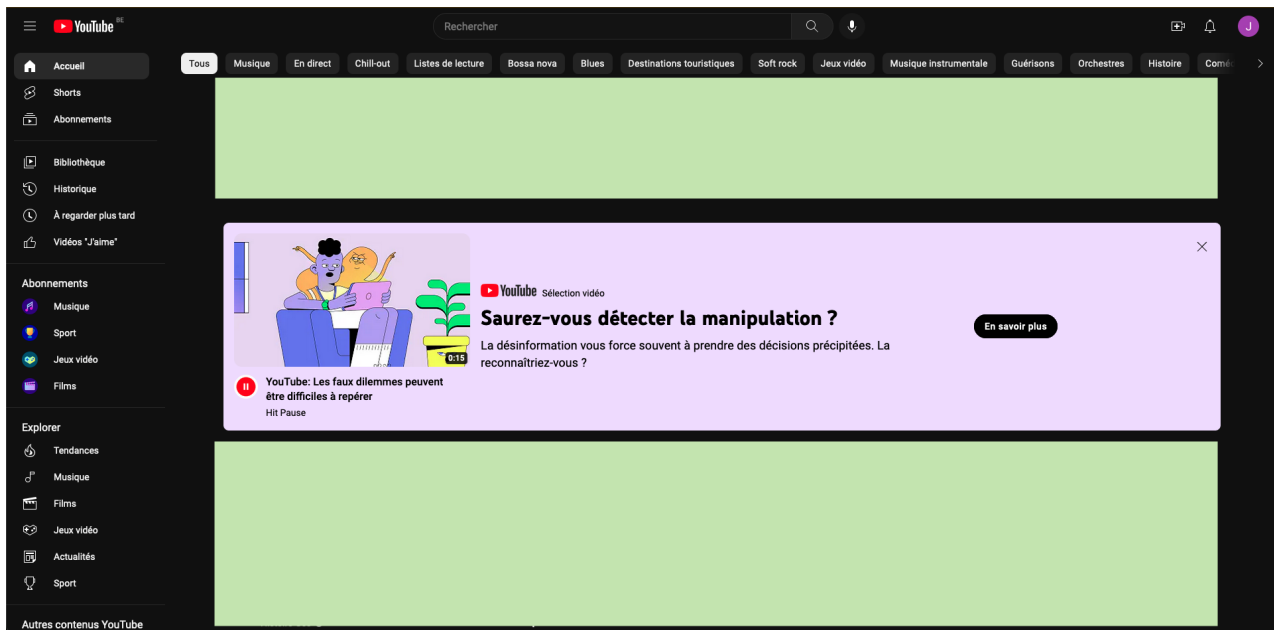


Figure E.16.: Date and time of the screenshot : Friday January 6, 2023 at 14:48

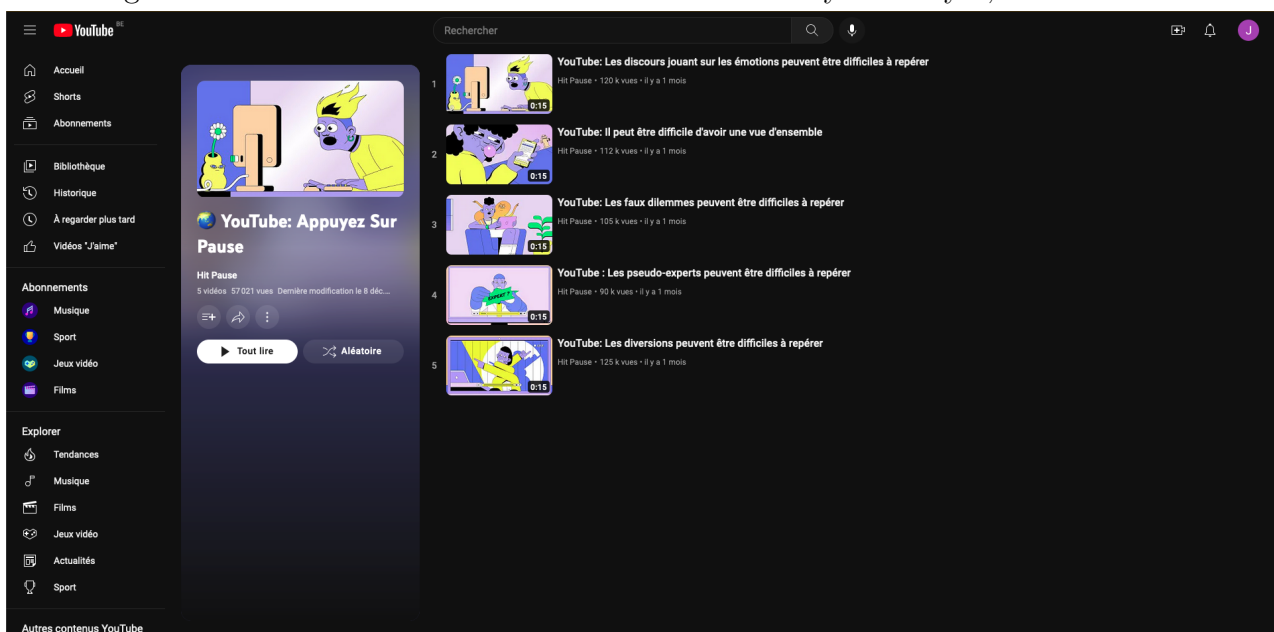


Figure E.17.: Date and time of the screenshot: Monday January 9, 2023 at 14:41

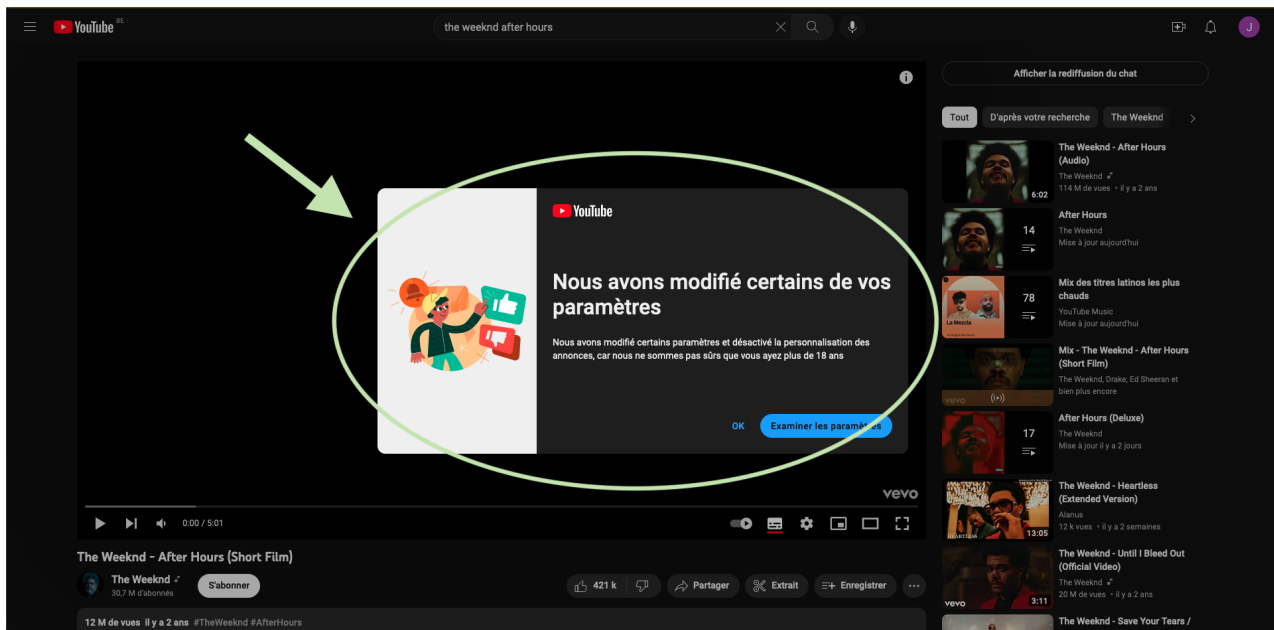


Figure E.18.: Date and time of the screenshot : Thursday January 12, 2023 at 20:26

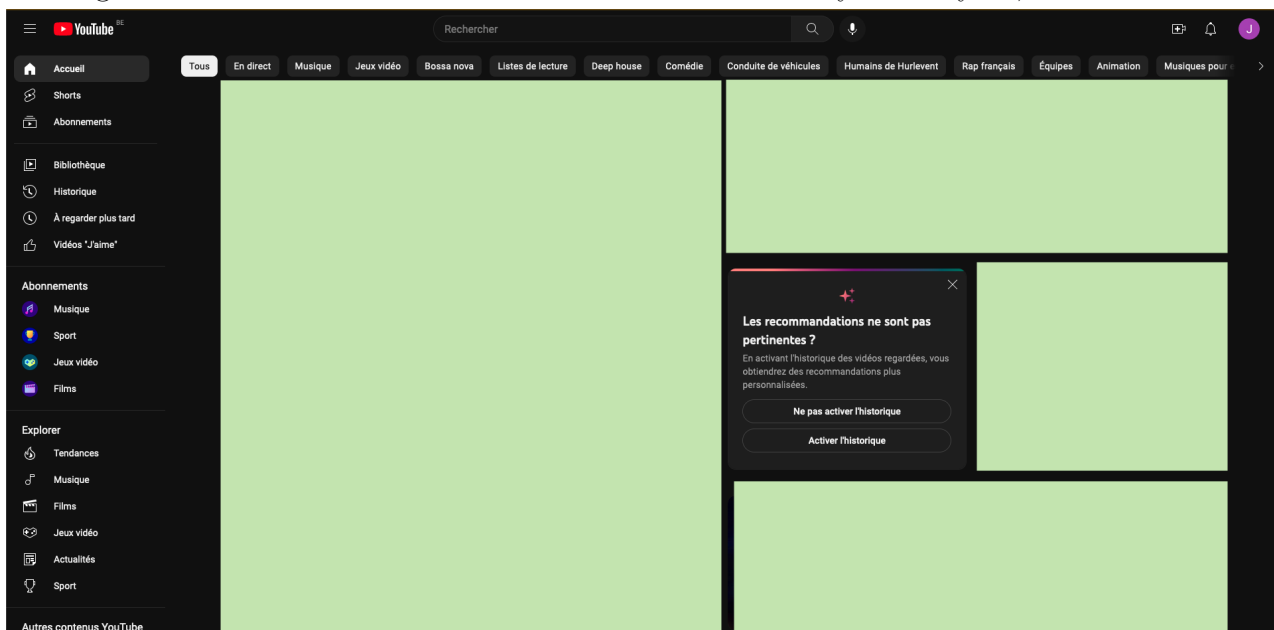


Figure E.19.: Date and time of the screenshot : Saturday 21 January 2023 at 18:41



Figure E.20.: Date and time of the screenshot : Sunday January 22, 2023 at 09:05

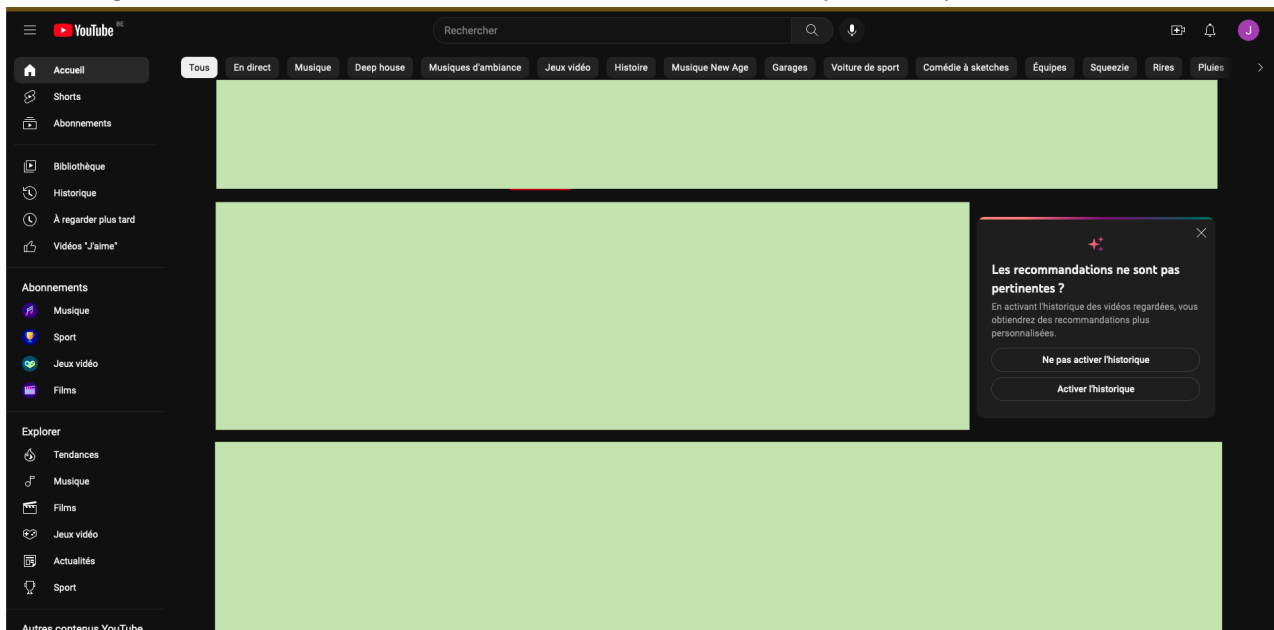


Figure E.21.: Date and time of the screenshot : Saturday 18 February 2023 at 13:35



Figure E.22.: Date and time of the screenshot : Friday March 3, 2023 at 09:15

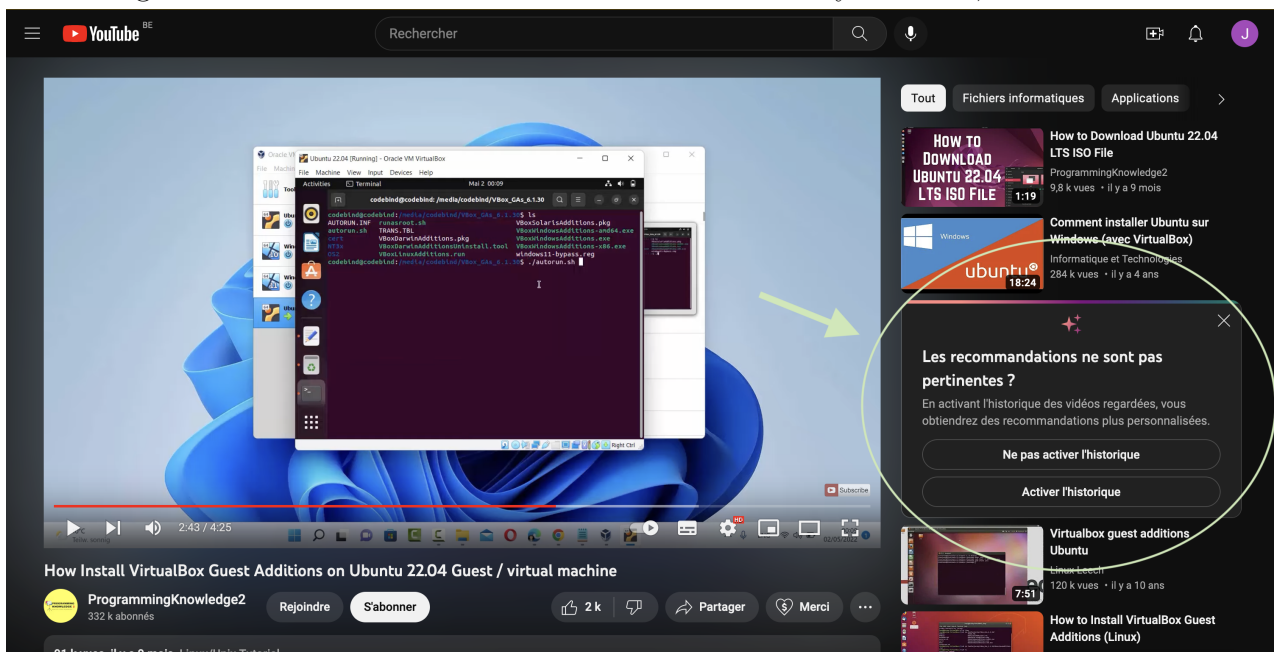
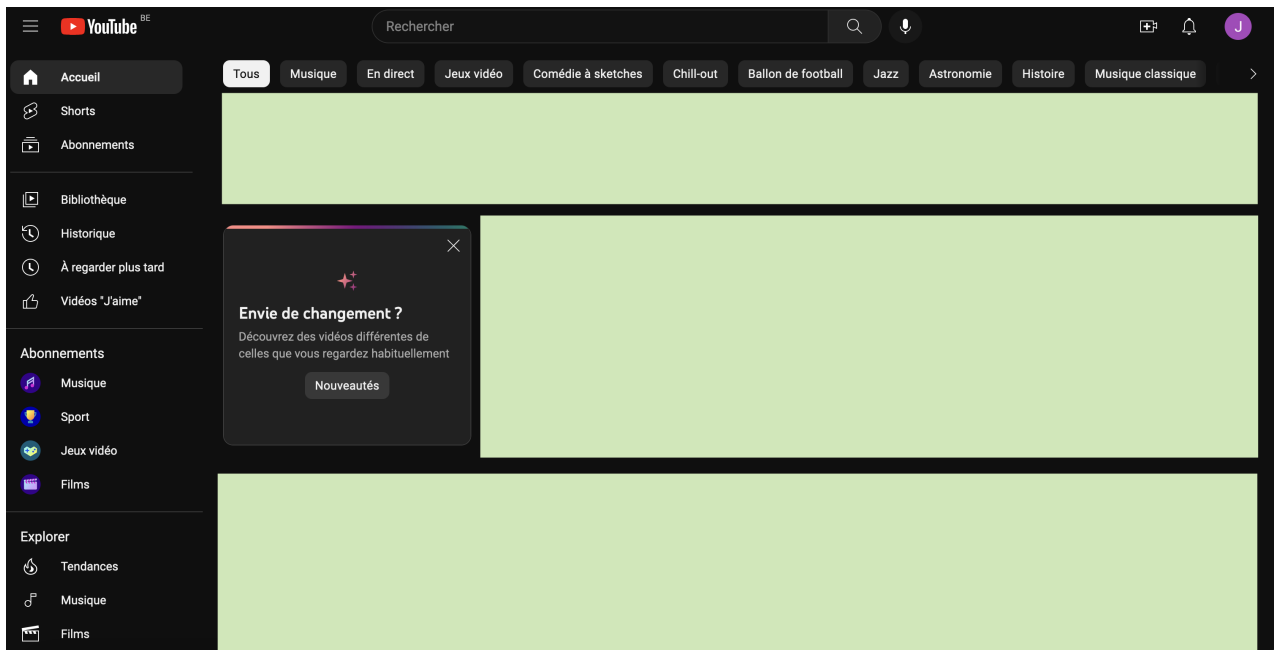


Figure E.23.: Date and time of the screenshot : Wednesday March 15, 2023 at 08:42



E.3.2. Prevention from Instagram

Figure E.24.: Instagram ad received on LinkedIn to control the time of use



Figure E.25.: Instagram ad received on LinkedIn to control the time of use



Figure E.26.: Instagram ad received on LinkedIn to control the time of use



E.3.3. Prevention from Facebook

Figure E.27.: New terms of use - Facebook (1)

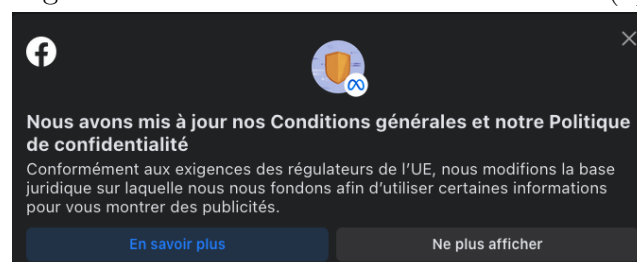


Figure E.28.: New terms of use - Facebook (2)

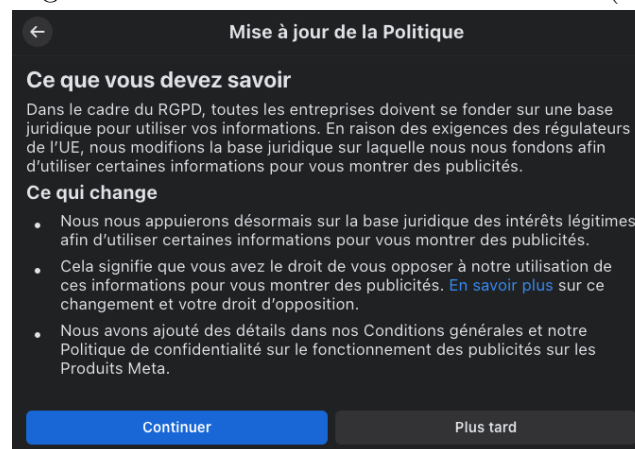
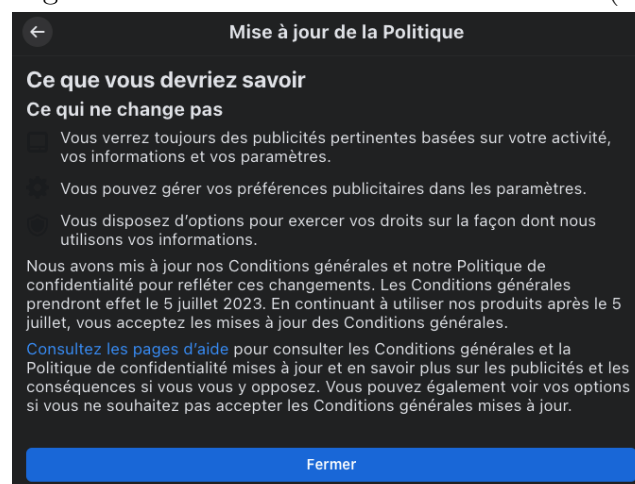


Figure E.29.: New terms of use - Facebook (3)



E.4. Youtube videos that talks about the impact of social networks on society

Note: in the following figures, although the videos are very interesting, some Youtubers do not quote their sources (like Lo Duff). It is therefore necessary to have a critical mind while looking at them.

Figure E.30.: Title of the video: This setting can break your smartphone addiction [138]. Link to the video: <https://www.youtube.com/watch?v=U0JUgXzC2CI&t=1s>

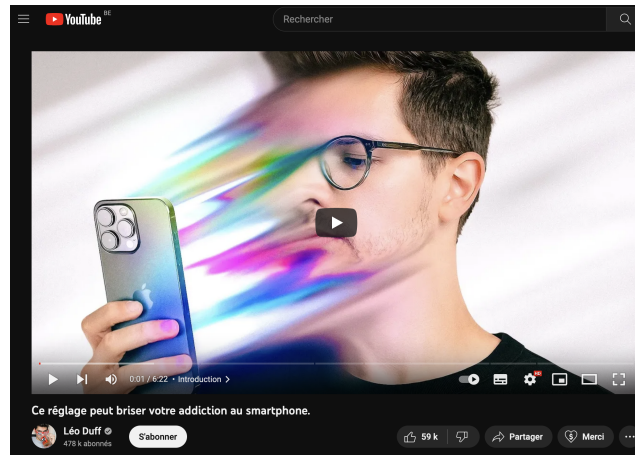


Figure E.31.: Title of the video: How TikTok has fucked up the brain of a generation [136]. Link to the video: <https://www.youtube.com/watch?v=FALDuZ5pPxc>

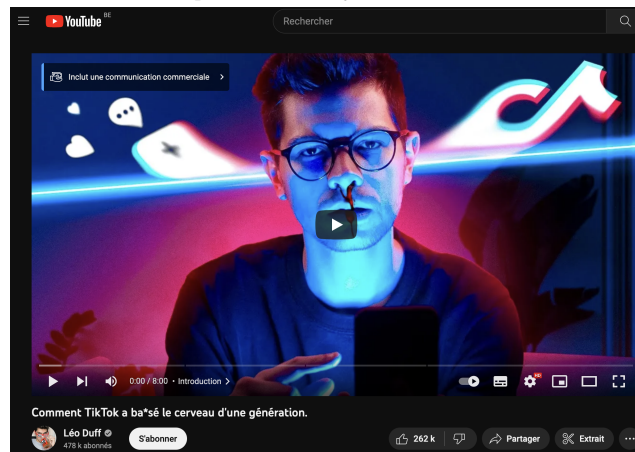


Figure E.32.: Title of the video: The system to reset your brain [137]. Link to the video: <https://www.youtube.com/watch?v=eUB6zkW-e5Y>

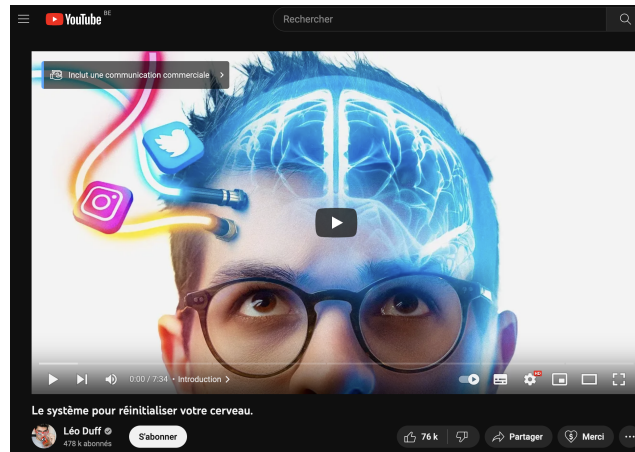


Figure E.33.: Title of the video: What Tinder doesn't tell you [165]. Link to the video: https://www.youtube.com/watch?v=bxQ_RngIQBI

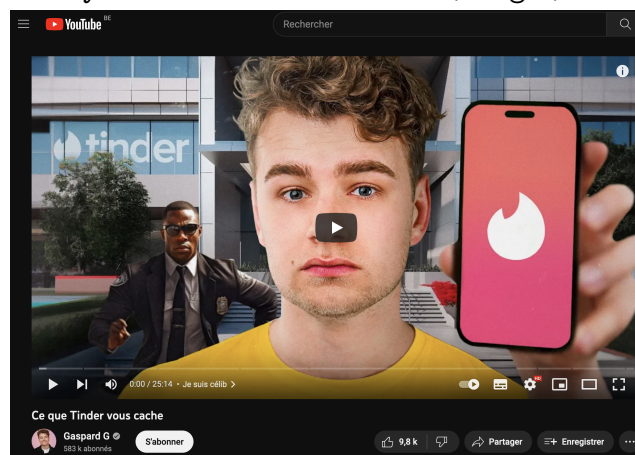


Figure E.34.: Title of the video: 300 Days of Watching TikTok Non-Stop: WHAT DOES IT DO? [158]. Link to the video: https://www.youtube.com/watch?v=LdNta5n_35Q

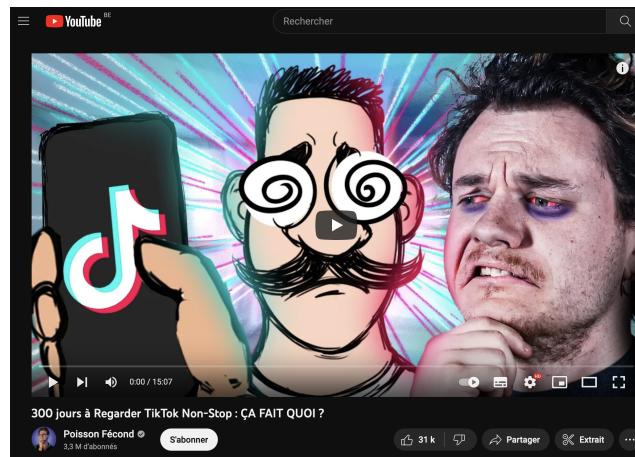


Figure E.35.: Title of the video: 50,000 Swipes on Tinder (and break his secret) [159]. Link to the video: <https://www.youtube.com/watch?v=9BJP2H1kZnk>

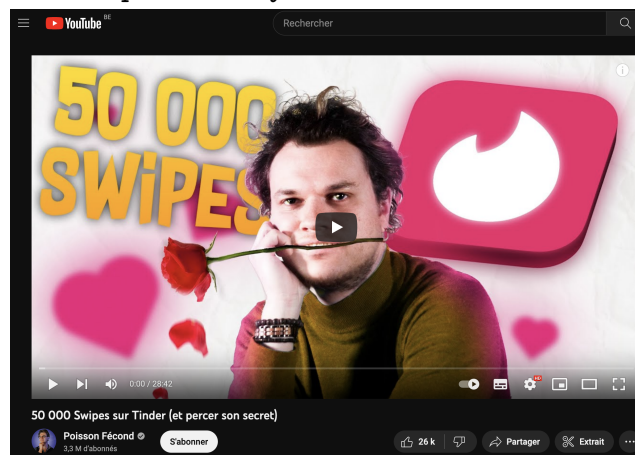
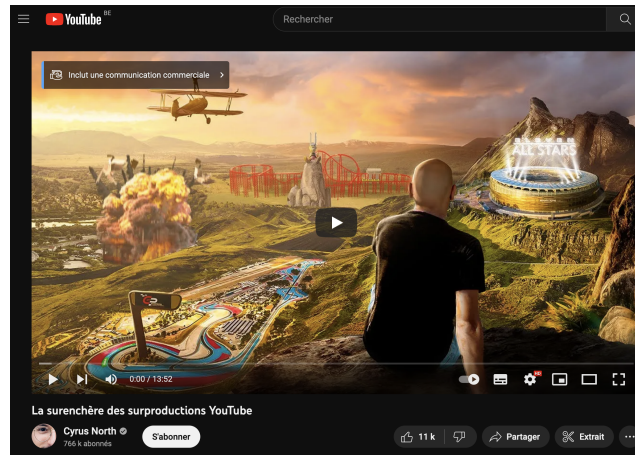


Figure E.36.: Title of the video: The overproduction of YouTube [275]. Link to the video: <https://www.youtube.com/watch?v=XgjyRVwHQQc>



UNIVERSITÉ CATHOLIQUE DE LOUVAIN

École polytechnique de Louvain

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve, Belgique | www.uclouvain.be/epl