

FREEDOM AND JUSTICE IN OUR TECHNOLOGICAL PREDICAMENT

HANS DE ZWART



Freedom and Justice in our Technological Predicament

Hans de Zwart

Master's Thesis

August 2018

Contents

Introduction	4
Part 1: What is going on?	6
The emerging logic of our digitizing world	6
From mediation	6
... to accumulation	7
... to centralization	13
Our technological predicament	14
Asymmetric relationships with arbitrary control	14
Data-driven appropriation	14
Domineering scale	15
Four Google case studies	16
Search	16
YouTube	19
Maps	21
reCAPTCHA	23
Part 2: How is that problematic?	26
Injustice in our technological predicament	26
The demands of justice as fairness	26
Lack of equality	28
Abuse of the commons	32
A utilitarian ethics	39
Unfreedom in our technological predicament	40
The demands of freedom as non-domination	41
The power to manipulate	45
Dependence on philanthropy	48
Arbitrary control	49
Part 3: What should we do about it?	53
Reducing the scale	53
Reinvigorating the commons	56
Equality in relationships	58
Acknowledgements	61
Bibliography	62

Introduction

As the director of an NGO advocating for digital rights I am acutely aware of the digital trail we leave behind in our daily lives. But I too am occasionally surprised when I am confronted with concrete examples of this trail. Like when I logged into my Vodafone account (my mobile telephony provider) and—buried deep down in the privacy settings—found a selected option that said: “Ik stel mijn geanonimiseerde netwerkgegevens beschikbaar voor analyse.”¹ I turned the option off and contacted Vodafone to ask them what was meant by anonymized network data analysis. They cordially hosted me at their Amsterdam offices and showed me how my movement behaviour was turned into a product by one of their joint ventures, Mezuro:

Smartphones communicate continuously with broadcasting masts in the vicinity. The billions of items of data provided by these interactions are anonymized and aggregated by the mobile network operator in its own IT environment and made available to Mezuro for processing and analysis. The result is information about mobility patterns of people, as a snapshot or trend analysis, in the form of a report or in an information system.²

TNO had certified this process and confirmed that privacy was assured: Mezuro has no access into the mobility information of individual people. From their website: “While knowledge of mobility patterns is of great social value, as far as we’re concerned it is certainly not more valuable than protecting the privacy of the individual.”³

Intuitively something about Vodafone’s behavior felt wrong to me, but I found it hard to articulate why what Vodafone was doing was problematic. This thesis is an attempt to find reasons and arguments that explain my growing sense of discomfort. It will show that Vodafone’s behavior is symptomatic for our current relationship with technology: it operates at a tremendous scale, it reuses data to turn it into new products and it feels empowered to do this without checking with their customers first.

The main research question of this thesis is how the most salient aspects of our technological predicament affect both justice as fairness and freedom as non-domination.

The research consists of three parts. In the first part I will look at the current situation to understand what is going on. By taking a closer look at the emerging logic of our digitizing society I will show how mediation, accumulation and centralization shape our technological predicament. This predicament turns out to be one where technology companies have a

¹Translated to English: “I make my anonymized network data available for analysis.”

²“Turning Big Data into Actionable Information.”

³“Understanding Mobility.”

domineering scale, where they employ a form of data-driven appropriation and where our relationship with the technology is asymmetrical and puts us at the receiving end of arbitrary control. A set of four case studies based on Google's products and services deepens and concretizes this understanding of our technological predicament.

In the second part of the thesis I will use the normative frameworks of John Rawls's justice as fairness and Philip Pettit's freedom as non-domination to problematize this technological predicament. I will show how data-driven appropriation leads to injustice through a lack of equality, the abuse of the commons, and a mistaken utilitarian ethics. And I will show how the domineering scale and our asymmetrical relationship to the technology sector leads to unfreedom through our increased vulnerability to manipulation, through our dependence on philanthropy, and through the arbitrary control that technology companies exert on us.

In the third and final part I will take a short and speculative look at what should be done to get us out of this technological predicament. Is it possible to reduce the scale at which technology operates? Can we reinvigorate the commons? And how should we build equality into our technology relationships?

Part 1: What is going on?

The digitization of our society is continuing with a rapid pace.⁴ The advent of the internet, and with it the World Wide Web, has been a catalyst for the transition from an economy which was based on dealing with the materiality of atoms towards one that is based on the immateriality of bits.

The emerging logic of our digitizing world

This digitization has made internet technology omnipresent in our daily lives. For example, 97.1% of Dutch people over 12 years old have access to the internet, 86.1% use it (nearly) every day and 79.2% access the internet using a smartphone (that was 40.3% in 2012).⁵ 12.1 million Dutch citizens have WhatsApp installed on their phone (that is around 90% of the smartphone owners) and 9.6 million people use the app daily. For the Facebook app these figures are 9.2 million and 7.1 million respectively.⁶ This turns out to have three main effects.

The digitization of society means that an increasing number of our interactions are technologically mediated. This mediation then enables a new logic of accumulation based on data. Together these two effects create a third: a centralizing force making the big become even bigger.

From mediation ...

It would be fair to say that many if not most of our interactions are technologically mediated⁷ and that we are all becoming increasingly dependent on internet based technologies.

This is happening with our social interactions, both in the relationships with our friends and in the relationships at work. Between two people speaking on the phone sits T-Mobile, between a person emailing their friends sits Gmail, to stay professionally connected we use

⁴“Our” is often an unspoken exclusive notion, so to make it explicit: This thesis is written from my perspective as a Dutch citizen. The concept of “our” and “we” in this thesis thus encompasses (parts of) society in North Western Europe. There are many parts of the world where the pace of digitization isn’t rapid and where the themes of this thesis will have very little bearing on daily reality.

⁵Data from 2017, see: “Internet: Toegang, Gebruik En Faciliteiten.”

⁶Data from June 2017, see: “Sociale Netwerken Dagelijks Gebruik Vs. App Geïnstalleerd Nederland.”

⁷In principle technology could have a very broad definition. You could argue that a book is a technology mediating between the reader and the writer. My definition of technology is a bit more narrow for this thesis. I am referring to the information and communication technologies that have accelerated the digitization of society and have categorically transformed it in the last thirty years or so (basically since the advent of the World Wide Web).

LinkedIn, and we reach out to each other using social media like Facebook, Twitter, Instagram and WhatsApp.

It is not just our social interactions which are mediated in this way. Many of our economic or commercial interactions have a third party in the middle too. We sell the stuff that we no longer want using online market places like eBay (and increasingly through social media like Facebook too), cash is slowly but surely being replaced by credit and debit cards and we shop online too. This means that companies like Amazon, Mastercard, and ING sit between us and the products we buy.

Even our cultural interactions are technologically mediated through the internet. Much of our watching of TV is done online, we read books via e-readers or listen to them as audio books, and our music listening is done via streaming services. This means that companies like YouTube, Netflix, Amazon, Audible, and Spotify sit between us and the cultural expressions and products of our society.

... to accumulation ...

This global architecture of networked mediation allows for a new logic of accumulation which Shoshana Zuboff—in her seminal article *Big Other*—calls “surveillance capitalism.”⁸ I will opt for the slightly more neutral term “accumulation”. Throughout her article, Zuboff uses Google as an example, basing a lot of her argument on two articles by Hal R. Varian, Google’s chief economist. According to Varian, computer mediated interaction facilitates new forms of contract, data extraction⁹ and analysis, controlled experimentation, and personalization and customization.¹⁰ These are the elements on which Google bases its playbook for business.

Conceptualizing the way that (our) data flows in these data economies, it is convenient to align with the phases of the big data life cycle. In *Big Data and Its Technical Challenges*, Jagadish et al. split the process up in: data acquisition; information extraction and cleaning; data integration, aggregation, and representation; modeling and analysis; and interpretation.¹¹ Many authors collapse these phases into a three-phase model: acquisition, analysis and application.¹² From the perspective of individual citizens or users this is a very clean way of looking at things: data is¹³ acquired, then something is done to it and finally it is

⁸Zuboff, “Big Other.”

⁹This is Varian’s euphemism for surveillance.

¹⁰Varian, “Computer Mediated Transactions,” 2.

¹¹Jagadish et al., “Big Data and Its Technical Challenges,” 88–90.

¹²See for example: Hirsch Ballin et al., “Big Data in Een Vrije En Veilige Samenleving,” 21.

¹³I will often use data with a singular verb, see: Rogers, “Data Are or Data Is?”

applied.¹⁴ Not all data flows in the same way in this accumulation ecosystem. It is therefore relevant to qualify the different ways in which this happens. For each of the phases, I will touch on some of the distinctions that can be made about the data.

Phase 1: Acquisition (gather as much data as possible)

Being the intermediary—the third party between users and the rest of their world—provides for a privileged position of surveillance. Because all the services of these intermediaries work with central servers, the accumulator can see everything their users do. It is trivial for Amazon to know how much time is spent on reading each book, which passages are the most highlighted, and what words are looked up the most in their dictionary. Amazon knows these things for each customer and at the aggregate level. Similarly, Spotify knows exactly what songs each individual likes to play most, and through that has a deep understanding about the current trends in music.

The costs (and size) of sensors is diminishing.¹⁵ This means that over the last couple of years it has become feasible to outfit users with products that have sensors (think microphones, cameras, GPS chips and gyro sensors). Every voice command that the user gives, every picture that is taken, and every route assisted trip is more data for the accumulator. These companies are now even starting to deliver sensors that go on (or in) our bodies, delivering data about sleep patterns, glucose levels, or general activity (like steps taken).

Some accumulators manage to get people to actually produce the data for them. Often this is data that can be turned into useful content for other users (like reviews of books on Amazon), or helps in solidifying trust in the network (reviews of Airbnb hosts and guests), and occasionally users are forced to provide data before they get access to a website (proving that you are human by clicking on photos).

Accumulators like Google and Facebook retain an enormous amount of data for each individual user,¹⁶ and even when they are forced to delete this personal data, they often resort to anonymization techniques in order to retain as much of the data as possible.¹⁷

Qualifying acquisition

The first distinction is whether the data relates to human beings at all. For most data that is captured via the internet or from our built environment this is the case, but there are

¹⁴This three-phase model also aligns with Zuboff's model of surveillance capitalism.

¹⁵Miller, "Cheaper Sensors Will Fuel The Age Of Smart Everything."

¹⁶Curran, "Are You Ready?"

¹⁷"How Google Retains Data We Collect."

domains where the data has nothing to do with us. It is assumed in what follows that we are talking about data that relates to humans.¹⁸

A dimension that will come back in all three phases is transparency. In this phase the question to ask is whether the person is aware that data is being collected and what data that is. This question can be asked for each individual, but it can also be asked in a more general way: is it possible to know what is being collected?

Another important distinction to make is whether the data is given voluntarily. Does the person have a choice about whether the data is given? This has an absolute side to it: is it possible for the person not to give this data? But more often there is some form of chained conditionality: given the fact that the person has decided to walk on this street, can they choose to not have their data collected? Has the person given their permission for the data to be acquired?

Often (but not always) related to this voluntariness is whether the data is collected as part of a private relationship between the person and the collector or whether the collection is done in the public sphere.

Furthermore, it is relevant to consider whether the data can be collected only once or whether it can be collected multiple times. A very similar question is whether it can only be collected by a single entity or whether others can collect it too.

Finally, it is worthwhile to think about whether the particular data is collected purposefully and with intent or whether the collection is a by-product of delivering another service.

Making the distinction between personal data (defined in Europe's General Data Protection Regulation as relating to an identified or identifiable individual¹⁹) and non-personal data probably isn't helpful in this phase. This data relates to human beings, and because it is very hard to anonymize data²⁰—requiring an active act by the collector of the data—it is probably best to consider all the data at this point in the process as personal data.

Phase 2: Analysis (use data scientists, experiments and machine learning to understand how the world works)

When you have a lot of data in a particular domain, you can start to model it to see how the domain works. If you collect a lot of movement data through people who use your mapping software to find their way, then you will gain a lot of insight into traffic patterns: Where is

¹⁸This isn't being too restrictive. As Karen Gregory writes: "Big data, like Soylent Green, is made of people." See: Gregory, "Big Data, Like Soylent Green, Is Made of People."

¹⁹"What Is Personal Data?"

²⁰Ohm, "Broken Promises of Privacy."

it busy at what time in the day? What happens to the traffic when a particular street gets broken up? If you also track news and events, then you would be able to correlate certain events (a concert in a stadium) with certain traffic patterns.

You no longer need to make an explicit model to see how the world works. Machine learning algorithms can use statistical methods to find correlational patterns. Chris Anderson (in)famously predicted that the tremendous amount of data that is being collected and available for analysis will make the standard scientific method—of making a hypothesis, creating a model, and finally testing the model—obsolete:

The new availability of huge amounts of data, along with the statistical tools to crunch these numbers, offers a whole new way of understanding the world. Correlation supersedes causation, and science can advance even without coherent models, unified theories, or really any mechanistic explanation at all.²¹

In certain domains it is possible to speed up the development of these machine learning algorithms by running experiments. If you have the ability to change the environment (hard to do with traffic, easy to do with a web interface), you can see how the behavior changes when the environment changes in a certain way. According to Anderson “Google and like-minded companies are sifting through the most measured age in history, treating this massive corpus as a laboratory of the human condition.”²²

Qualifying analysis

There are basically three possible results when it comes to a person’s data at the end of this phase:

1. The person is still identifiable as that person.
2. The data is pseudonymized or anonymized, but is still individualized. There is still a direct relationship between the collected data from the person and how it is stored.
3. The data is aggregated into some form that no longer relates to an individual. The person has become part of a statistic or a weight in some probabilistic model.

Of course it can also be a combination of these three results. They are in no way mutually exclusive.

Once again, it might be also be a relevant distinction to see how transparent it is for the person as to how their data is being stored.

²¹Anderson, “The End of Theory.”

²²Ibid.

Phase 3: Application (use the model to create predictions and sell these)

When you understand how a particular domain works, you can use that understanding to predict the future. If you know the current circumstances and you know what usually happens in these circumstances, you can start to make predictions and sell them on the market.

The dominant market for predictions at this point in time is advertising. Companies like Google and Facebook use this logic of accumulation to try and understand buying intent and sell this knowledge as profiles to advertise against. Facebook for example, allows you to target on demographics, location, interests (“Find people based on what they’re into, such as hobbies, favourite entertainment and more.”) and behaviours (“Reach people based on their purchasing behaviours, device usage and other activities”).²³ Some marketeers have gone through the trouble to list out all of Facebook’s ad targeting options.²⁴ These include options like “Net Worth over \$2,000,000”, “Veterans in home”, “Close Friends of Women with a Birthday in 0-7 days”, “Likely to engage with political content (conservative)”, “Active credit card user”, “Owns: iPhone 7” and “African American (US) Multicultural Affinity.”²⁵ It is important to note that many of these categories are not based on data that the user has explicitly or knowingly provided, but are based on a calculated probability instead.

Advertising isn’t the only market where predictions can be monetized, the possibilities are endless. Software predicts crime and sells these predictions to the police,²⁶ software predicts the best performing exchange-traded funds in each asset class and sells these predictions as automatic portfolio management to investors,²⁷ and software predicts which patients will most urgently need medical care and sells these predictions to hospitals.²⁸ Some people label this moment in time as “the predictive turn”.²⁹

Qualifying application

This is the phase where the data that has been acquired and analyzed is used back into the world. The first relevant distinction is whether the use of the data (directly) affects the person from which the data was acquired. Is there a direct relationship?

Next, it is important to look at whether the data is applied in the same domain (or within the same service) as where it was acquired. Or is it acquired in one domain and then used

²³“Choose Your Audience.”

²⁴“Facebook Ads Targeting Comprehensive List.”

²⁵I find this final category deeply problematic, see: De Zwart, “Facebook Is Gemaakt Voor Etnisch Profileren.”

²⁶“Predictive Policing Software.”

²⁷“Wealthfront Investment Methodology White Paper.”

²⁸“DeepMind Health.” DeepMind’s slogan on their homepage is “Solve intelligence. Use it to make the world a better place.”

²⁹Van Hoboken, “Comment on ‘Democracy Under Siege’.”

in another? If that is the case, then often the application itself is part of the acquisition of data in some other process.

The distinction between private use or public use of the data is interesting too. Sometimes this distinction is hard to make, so the cleanest way to draw the line is between proprietary data and data that can be freely used and shared. Another way of exploring the distinction between private and public use is to ask where (the majority of) the value accrues. Closely related to this point is the question of whether the use of the data aligns with what the person finds important. Is the use of the data (socially) beneficial from the person's perspective?

Of course it is again relevant whether it is transparent to the person how the data is applied.

Data appropriation

Having looked at the three phases of accumulation it becomes possible to create a working definition of data appropriation. To “appropriate” in regular use, means to take something for one's own use, typically without the owners permission³⁰ or to take or make use of without authority or right.³¹ The definition of “data appropriation” can stay relatively close to that meaning. Data is appropriated from a person when all three of the following conditions are true:

1. The data originates with that person.
2. The organization that acquires, analyses or applies the data isn't required by law to collect, store, or use the data.
3. Any one of the following conditions is true:
 - The data is acquired against their volition (i.e. involuntarily).
 - The data is acquired without their knowledge.
 - The data is applied against their volition.
 - The data is applied without their knowledge.

It is important to note that what is done to the data in the analysis phase—whether the data is pseudonymized, anonymized or used at an aggregate level—has no bearing on whether the use is to be considered as appropriative. So the fact that there might not have been a breach of privacy (or of contextual integrity) does not mean there was no appropriation. And similarly, it doesn't matter for what purposes the data is applied. Even if the application can only serve towards a public social benefit, it might still have been appropriation that enabled the application.

³⁰“Definition of Appropriate in the Oxford Dictionary.”

³¹“Definition of Appropriate in the Merriam Webster Dictionary.”

... to centralization

Mediation and accumulation create a third effect: they lead to centralization. Initially we thought that the internet would be a major source of disintermediation and would remove the intermediaries from our transactions. Robert Gellman's 1996 article *Disintermediation and the Internet* is illustrative of this idea. He wrote:

The Internet offers easy, anonymous access to any type of information product [...]. The traditional intermediaries—newsstands, book stores, and video stores—are not necessary. [...] With the Internet the traditional intermediaries are swept away. Anyone of any age who can click a mouse can access any public server on the network. The limitations that were inherent in traditional distribution methods are no longer there.³²

Allowing for direct (often even peer-to-peer) connections, we would be able to decrease our dependence on companies earning their money through offering different options to their customers. We no longer needed travel agents to book holidays, or real estate agents to buy and sell houses. And news would find us directly rather than having to be bundled into a newspaper.

A more truthful description of what turned out to be happening is that we switched out one type of intermediary for another. Rather than being dependent on travel agents, realtors and newspapers, we became dependent on companies like Google, Facebook, and Amazon. According to Ben Thompson, to be successful in the pre-internet era you either had to have a monopoly or you needed to control distribution. The internet has changed this. Distribution of digital goods is free and transaction costs are zero (meaning you can scale to billions of customers):

Suppliers can be aggregated at scale leaving consumers/users as a first order priority. [...] This means that the most important factor determining success is the user experience: the best distributors/aggregators/market-makers win by providing the best experience, which earns them the most consumers/users, which attracts the most suppliers, which enhances the user experience in a virtuous cycle.³³

Thompson calls this “aggregation theory”, and uses it to explain the success of Google's search, Facebook's content, Amazon's retail goods, Netflix's and YouTube's videos, Uber's drivers, and Airbnb's rooms. Aggregation theory has a centralizing effect:

³²Gellman, “Disintermediation and the Internet,” 7.

³³Thompson, “Aggregation Theory.”

Thanks to these virtuous cycles, the big get bigger; indeed, all things being equal the equilibrium state in a market covered by Aggregation Theory is monopoly: one aggregator that has captured all of the consumers and all of the suppliers.³⁴

It is interesting to note that the aggregators don't create their monopoly by limiting the options the internet user has. It could even be said that the user chooses to be inside the aggregator's monopoly because of the better user experience.³⁵ However, it is the monopolist which in the end has the singular ability to fully shape the user's experience.

Our technological predicament

It is now clear how mediation allows for a new logic of accumulation which then keeps on accelerating through centralization. Each of these effects results in a particular salient characteristic of our technological predicament. Mediation leads to asymmetric relationships with arbitrary control, accumulation leads to data-driven appropriation, and centralization leads to a domineering scale.

Asymmetric relationships with arbitrary control

The relationship between technology companies and their users is one where the former can afford to make unilateral and completely arbitrary decisions. It is the company that decides to change the way a product looks or works, and it is the company that can decide to give the user access or to block their account. This leads to a loss of control (the company making the choices instead of the user), often with few if any forms of redress in case something happens that the user doesn't like.

There is also a clear asymmetry in transparency. These companies have a deep knowledge about their users, and the users most times can only know very little about the company.

Data-driven appropriation

The technology companies base their services—and get their quality from—the data that they use as their input. Often this data is given by the user through using the product or through giving their attention, sometimes the user is actively turned into a data collector,

³⁴Thompson, "Antitrust and Aggregation."

³⁵This is also one of the reasons why classical antitrust thinking doesn't have the toolkit to address this situation.

and occasionally these companies are free-riders on other services that are open enough to allow them to use their data.

It is important to accentuate the nontransparent nature of much of what these companies do. Often the only way to try to understand the way they use data, is through a black box methodology, trying to see what goes into them and what comes out of them, and using that information to try and piece together the whole puzzle. The average user will have little insight or knowledge about how the products they use every day work, or what their larger impact might be.

Even if there is the option not to share your data with these companies, then there still is what Solon Barocas and Helen Nissenbaum call the tyranny of the minority: “The willingness of a few individuals to disclose information about themselves may implicate others who happen to share the more easily observable traits that correlate with the traits disclosed.”³⁶

Technology companies have the near classic feature of capitalism: they manage to externalize most of the costs and the negative societal consequences that are associated with the use of their products, while also managing to hold on to a disproportionate amount of the benefits that accrue. These costs that have to be borne by society aren’t spread out evenly. The externalities have disparate impacts, usually strengthening existing divisions of power and wealth.

Domineering scale

Centralization is the reason why these technology companies can operate at a tremendous scale. Their audience is the (connected) world and that means that a lot of what they do results in billions of interactions. The technology giants that are so central to our lives mostly have a completely dominant position for their services or products. In many cases they have a de facto monopoly, with the accompanying high level of dependence for its users.

The fact that information based companies have very recently replaced oil companies in the charts listing the largest companies in the world by market value³⁷ is clear evidence of this dominance.

³⁶Barocas and Nissenbaum, “Big Data’s End Run Around Procedural Privacy Protections,” 32.

³⁷The top ten at the end of the first quarter of 2011 were Exxon Mobil, PetroChina, Apple Inc., ICBC, Petrobras, BHP Billiton, China Construction Bank, Royal Dutch Shell, Chevron Corporation, and Microsoft. At the end of the 1st quarter of 2018, Apple Inc., Alphabet Inc., Microsoft, Amazon.com, Tencent, Berkshire Hathaway, Alibaba Group, Facebook, JPMorgan Chase en Johnson & Johnson were at the top of list. See: “List of Public Corporations by Market Capitalization.”

Four Google case studies

So far, the discussion about our technological predicament has stayed at an abstract level. I will use a set of four case studies to make our predicament more concrete and both broaden the conceptions about what can be done through accumulated data, and deepen the understanding about how that is done.

All of these case studies are taken from the consumer product and services portfolio of Google,³⁸ as one of the world's foremost accumulators. Most readers will be familiar with—and users of—these services. I want to highlight some of the lesser known aspects of these products and show how all of them have the characteristics of asymmetrical relationships, data-driven appropriation and a domineering scale. Although the selection of these particular cases is relatively arbitrary,³⁹ together they do span the territory of practices that will turn out to be problematic in the second part of this thesis.

Search

Google is completely dominating the search engine market. Worldwide—averaging over all devices—their market share is about 75%.⁴⁰ But in certain markets and for certain devices this percentage is much higher, often above 90%.⁴¹ Every single one of the more than 3.5 billion daily searches⁴² is used by Google to further tweak its algorithms and make sure that people find what they are looking for. Search volume drives search quality,⁴³ and anybody who has ever tried any other search engine knows that Google delivers the best results by far.⁴⁴

A glance at anyone's search history will show that Google's search engine is both used to look up factual information (basically it is a history of things that this person didn't know yet), as well as the transactional intentionality of that user (what that person is intending to do, buy, or go to). On the basis of this information, Google is able to infer many things about this person, including for example what illnesses the person might have (or at least their symptoms), what they will likely vote at the next election, and what their job is. Google

³⁸Google is now a wholly owned subsidiary of Alphabet, but all these examples still fall under the Google umbrella. See: Page, "G Is for Google."

³⁹Alphabet literally has products starting with every letter of the alphabet. See: Murphy and Rathi, "All of Google's—er, Alphabet's—companies and Products from A to Z."

⁴⁰"Search Engine Market Share."

⁴¹In the Netherlands for example, Google Search has a 89% market share on the desktop and a 99% market share on mobile. See: Borgers, "Marktaandeelen Zoekmachines Q1 2018."

⁴²"Google Search Statistics."

⁴³Levy, "How Google's Algorithm Rules the Web."

⁴⁴For most use cases, there are specific domains where niche search engines might perform better.

even knows when this person is sleeping, as these are the moments when that person isn't doing any searching.

Google makes some of its aggregated search history available for research through Google Trends.⁴⁵ You can use this tool to look up how often a particular search is done. Google delivers this data anonymously, you can't see who is searching for what. In his book *Everybody Lies*, Seth Stephens-Davidowitz has shown how much understanding of the world can be gleaned through this tool. He contends that Google's search history is a more truthful reflection of what people think than any other way of assessing people's feelings and thoughts. People tell their search engines things they wouldn't say out in the open. Unlike our behavior on social media like Facebook, we don't only show our good side to Google's search engine. Stephens-Davidowitz became famous for his research using the percentage of Google search queries that include racist language to show that racism is way more prevalent in the United States than most surveys say it is. He used Google data to make it clear that Obama lost about 4 percentage points in the 2008 vote, just because he was black.⁴⁶ Stephens-Davidowitz is "now convinced that Google searches are the most important dataset ever collected on the human psyche."⁴⁷ We shouldn't forget that he was able to do his research though looking at Google search history as an outsider, in a way reverse engineering the black box.⁴⁸ Imagine how much easier it would be to do this type of research from the inside.

It often feels like Google's search results are a neutral representation of the World Wide Web, algorithmically surfacing what is most likely to be the most useful information to deliver as the results for each search, and reflecting what is searched by the searching public at large. But it is important to realize two things: Firstly, what Google says about you, frames to a large extent how people see you. And secondly, the search results are not neutral, but are a reflection of many of society's biases.

The first page of search results when you do a Google search for your full name, in combination with the way how Google presents these results (do they include pictures, videos, some snippets of information), have a large influence on how people initially see you. This is even more true in professional situations and in the online space. You have very little influence about what information is shown about you on this first page.

This fact is the basis of the now famous case at the European Court of Justice, pitting Google against Mario Costeja González and the Spanish Data Protection Authority. Costeja

⁴⁵"Google Trends."

⁴⁶Stephens-Davidowitz, "The Cost of Racial Animus on a Black Candidate," 36.

⁴⁷Stephens-Davidowitz, *Everybody Lies*, 14.

⁴⁸Google noticed Stephens-Davidowitz's research and hired him as a data scientist. He stayed on for one and a half years. See: "About Seth."

González was dismayed at the fact that a more than ten year old piece of information, from a required ad in a newspaper, describing his financial insolvency, was still ranking high in the Google search results for his name, even though the information was no longer directly relevant. The court realized the special nature of search engine results:

Since the inclusion in the list of results, displayed following a search made on the basis of a person's name, of a web page and of the information contained on it relating to that person makes access to that information appreciably easier for any internet user making a search in respect of the person concerned and may play a decisive role in the dissemination of that information, it is liable to constitute a more significant interference with the data subject's fundamental right to privacy than the publication on the web page.⁴⁹

The Court told Google to remove the result at the request of Costeja González. This allowed him to exercise what came to be called “the right to be forgotten”, but what should really be called “the right to be delinked”. In her talk *Our Naked Selves as Data – Gender and Consent in Search Engines*, human rights lawyer Gisela Perez de Acha talks about her despair at Google still showing the pictures of her topless FEMEN-affiliated protest from a few years back. Google has surfaced her protest as the first thing people see about her when you look up her name. In the talk, she wonders what we can do to fight back against private companies deciding who we are online.⁵⁰

That Google's search results aren't neutral, but a reflection of society's biases, is described extensively by Safiya Umoja Noble in her book *Algorithms of Oppression*. The starting point for Noble is one particular moment in 2010:

While Googling things on the Internet that might be interesting to my step-daughter and nieces, I was overtaken by the results. My search on the keywords “black girls” yielded HotBlackPussy.com as the first hit.⁵¹

For Noble this is a reflection of the way that black girls are hypersexualized in American society in general. She argues that advertising in relation to black girls is pornified and that this translates itself into what Google decided to show for these particular keywords. The reflection of this societal bias can be found in many more examples of search results. For example when searching for “three black teenagers” (showing inmates)⁵² or “unprofessional hairstyles for work” (showing black women with natural hair that isn't straightened).⁵³

⁴⁹Court of Justice, “Google Spain SL and Google Inc. V Agencia Española de Protección de Datos (AEPD) and Mario Costeja González,” para. 87.

⁵⁰Perez de Acha, “Our Naked Selves as Data – Gender and Consent in Search Engines.”

⁵¹Noble, *Algorithms of Oppression*, 3.

⁵²Alli, “YOOWWWW LOOK AT THIS.”

⁵³Kamona, “I Saw a Tweet Saying ‘Google Unprofessional Hairstyles for Work’.”

The lack of a black workforce at Google,⁵⁴ and the little attention that is paid to the social sciences in the majority of engineering curriculums, don't help in raising awareness and proactively preventing the reification of these biases. Usually Google calls these results anomalies and beyond their control. But Noble asks: "If Google isn't responsible for its algorithm, then who is?"⁵⁵

YouTube

YouTube is the second largest search engine in the world (after Google's main search engine).⁵⁶ More than 400 hours of video are uploaded to YouTube every minute,⁵⁷ and together we watch more than a billion hours of YouTube videos every single day.⁵⁸⁵⁹ It is safe to say that YouTube is playing a very big role in our lives.

I want to highlight three central aspects about YouTube. First, I will show how Google regulates a lot of our expression through the relatively arbitrary blocking of YouTube accounts. Next, I will show how the data-driven business model, in combination with the ubiquity and commodification of artificial intelligence, leads to some very surprising results. Finally, I will show how Google relies on the use of free human labor to increase the quality of its algorithmic machine.

Women on Waves is an organization which "aims to prevent unsafe abortions and empower women to exercise their human rights to physical and mental autonomy."⁶⁰ It does this through providing abortion services on a ship in international waters. In recent years, they've also focused on providing women internationally with abortion pills, so that they can do medical abortions. Women on Waves has YouTube videos in many different languages showing how to do this safely.⁶¹ In January 2018, their YouTube account was suspended for violating what YouTube calls its "community guidelines". Appeals through the appeals process didn't help. After creating some negative media attention about the story, their account got reinstated and Google issued a non-apology for an erroneous block. Unfortunately since then a similar suspension happened at least two more times with similar

⁵⁴In 2017, the percentage of black tech workers at Google was 1.4%. See: "Annual Report - Google Diversity."

⁵⁵Noble, *Algorithms of Oppression*, 80.

⁵⁶Smith, "39 Fascinating and Incredible YouTube Statistics."

⁵⁷Brouwer, "YouTube Now Gets Over 400 Hours Of Content Uploaded Every Minute."

⁵⁸Goodrow, "You Know What's Cool?"

⁵⁹This last figure is particularly staggering. It means that if you look up any world citizen at any point in time, the chances that they are watching a YouTube video right when you drop in, is bigger than 1 in 200. Or said in another way: Globally we spend more than 0.5% of the total time that we have available to us watching videos on YouTube.

⁶⁰"Who Are We?"

⁶¹Being present on YouTube is important for them because in many countries it is safer to visit youtube.com than womenonwaves.org.

results. YouTube refuses to say why and how these blocks happen (hiding behind “internal information”), and says that they have to take down so much content every day that mistakes are bound to be made.⁶² This is of course just one example of legion. According to Evelyn Austin, the net result of this situation is that “users have become passive participants in a Russian Roulette-like game of content moderation.”⁶³

Late 2017, artist James Bridle wrote a long essay about the near symbiotic relationship between younger children and YouTube.⁶⁴ According to Bridle, children are often mesmerized by a diverse set of YouTube videos: from nursery rhymes with bright colours and soothing sounds to surprise egg unboxing videos. If you are a YouTube broadcaster and want to get children’s attention (and the accompanying advertising revenue), then one strategy is to copy and pirate other existing content. A simple search for something like “Peppa Pig”, gives you results where it isn’t completely obvious which are the real videos and which are the copies. Branded content usually functions as a trusted source. But as Bridle writes:

This no longer applies when brand and content are disassociated by the platform, and so known and trusted content provides a seamless gateway to unverified and potentially harmful content.⁶⁵

YouTube creators also crank up their views through using the right keywords in the title. So as soon as something is popular with children, millions of similar videos will be created, often by bots. Bridle finds it hard to assess the degree of automation, as it is also often real people acting out keyword driven video themes. The vastness of the system, and the many languages in which these videos are available, creates a dimensionality that makes it hard to think about and understand what is actually going on. Bridle makes a convincing point, that for many of these videos neither the creator or the distribution platform has any idea of what is happening. He then goes on to highlight the vast number of videos that use similar tropes, but contain a lot of violence and abusive scenes. He can’t find out who makes them and with what intention, but it is clear that they are “feeding upon a system which was consciously intended to show videos to children for a profit” and for Bridle it is also clear that the “system is complicit in the abuse.”⁶⁶ He thinks YouTube has a responsibility to deal with this, but can’t really see a solution other than dismantling the system. The scale is too big for human oversight and there is no nonhuman oversight which can adequately address the situation that Bridle has described. To be clear, this is not just about children videos. It would be just as easy to write a completely similar narrative about “white nationalism,

⁶²Austin, “Women on Waves’ Three YouTube Suspensions This Year Show yet Again That We Can’t Let Internet Companies Police Our Speech.”

⁶³Ibid.

⁶⁴Bridle, “Something Is Wrong on the Internet.”

⁶⁵Ibid.

⁶⁶Ibid.

about violent religious ideologies, about fake news, about climate denialism, about 9/11 conspiracies.”⁶⁷

The conspiratorial nature of many of the videos on YouTube is problematic for the platform. It therefore announced in March 2018, that they would start posting information cues to fact-based content alongside conspiracy theory videos. YouTube will rely on Wikipedia to provide this factual information.⁶⁸ It made this announcement without consulting with Wikimedia, the foundation behind Wikipedia. As Louise Matsakis writes in Wired:

YouTube, a multibillion-dollar corporation flush with advertising cash, had chosen to offload its misinformation problem in part to a volunteer, nonprofit encyclopedia without informing it first.⁶⁹

Wikipedia exists because millions of people donate money to the foundation and because writers volunteer their time into making the site into what it is. Thousands of editors monitor the changing contents of the encyclopedia, and in particular the pages that track conspiracy theories usually have years of active work inside of them.⁷⁰ YouTube apparently had not considered what impact the linking from YouTube would have on Wikipedia. This is not just about the technological question of whether their infrastructure could handle the extra traffic, but also what it would do to the editor community if the linking would lead to extra vandalism for example. Wikipedian Phoebe Ayers tweeted: “It’s not polite to treat Wikipedia like an endlessly renewable resource with infinite free labor; what’s the impact?”⁷¹

Maps

Whenever I do a presentation somewhere in the Netherlands, I always ask people to raise their hand if they have used Google Maps to reach the venue. Most times a large majority of the people have done exactly that. It has become so ubiquitous that it is hard to imagine how we got to where we needed to be, before it existed. The tool works so well that most people will just blindly follow its instructions most of the time. Google Maps is literally deciding what route we take from the station to the theatre.

Even though maps are highly contentious and deeply political by nature,⁷² we still assume

⁶⁷Ibid.

⁶⁸Matsakis, “YouTube Will Link Directly to Wikipedia to Fight Conspiracy Theories.”

⁶⁹Matsakis, “Don’t Ask Wikipedia to Cure the Internet.”

⁷⁰Farokhmanesh, “YouTube Didn’t Tell Wikipedia About Its Plans for Wikipedia.”

⁷¹Ayers, “YouTube Should Probably Run Some A/B Tests with the Crew at @WikiResearch First.”

⁷²Google follows local laws when presenting a border, so when you look up the Crimea from the Russian version of Google Maps you see it as part of Russia, whereas if you look at it from the rest of the world it will be listed as disputed territory. See: Chappell, “Google Maps Displays Crimean Border Differently In Russia,

that they are authoritative and in some way neutral. I started doubting this for the first time when I found out that Google Maps would never route my cycle rides through the canals of Amsterdam, but would always route me around them, even if this was obviously slower.⁷³ One of my friends was sure that rich people living on the canals had struck a deal with Google to decrease the traffic in front of their house. I attributed it to Google's algorithms being more attuned to the street plan of San Francisco than to those of a World Heritage site designed in the 17th century.

But then I encountered the story of the Los Angeles residents living at the foot of the hills that harbor the Hollywood Sign. They've been on a mission in the past couple of years to wipe the Hollywood Sign of the virtual map, because they don't like tourists parking in the streets.⁷⁴ And for a while they were successful: when you were at the bottom of the hill and asked Google Maps for a route, the service would tell you to walk for one and a half hours to a viewing point at the other end of the valley, instead of showing that a walking path exists that will take you up the hill in 15 minutes. This tweak to the mapping algorithm is just one of the countless examples of where Google applies human intervention to improve their maps.⁷⁵ As users of the service, we can't see how much human effort has gone into tweaking the maps to give the best possible results. This is because (as I wrote in 2015) "every design decision, is completely mystified by a sleek and clean interface that we assume to be neutral."

Next to human intervention, Google also uses algorithms based on artificial intelligence to improve the map. The interesting thing about internet connected digital maps is that they allow for the map to change on the basis of what their users are doing. Your phone and all the other phones on the road are constantly communicating with Google's services, and this makes it possible for Google to tell you quite precisely when you are going to hit a traffic jam. In 2016, Google rolled out an update to its maps to highlight in orange "areas of interest [...], places where there's a lot of activities and things to do."⁷⁶ Google decides on which areas are of interest through an algorithm (with the occasional human touch in high-density areas): "We determine 'areas of interest' with an algorithmic process that allows us to highlight the areas with the highest concentration of restaurants, bars and shops."⁷⁷

This obviously begs the question: interesting for whom? Laura Bliss found out that the ser-

U.S."

⁷³I've written up this example before. See: De Zwart, "Demystifying the Algorithm." and De Zwart, "Google Wijst Me de Weg, Maar Niet Altijd de Kortste."

⁷⁴The residents argue that fire trucks aren't able to pass by these parked cars in case of an emergency.

⁷⁵The project to improve the quality of the maps at Google is called 'Ground Truth'. See: Madrigal, "How Google Builds Its Maps—and What It Means for the Future of Everything."

⁷⁶Li and Bailang, "Discover the Action Around You with the Updated Google Maps."

⁷⁷Ibid.

vice didn't highlight streets that were packed with restaurants, businesses and schools in relatively low-income and predominantly non-white areas. Real life divides are now manifested in a new way according to Bliss. She asks the largely rhetorical questions: "Could it be that income, ethnicity, and Internet access track with 'areas of interest' " with the map literally highlighting the socio-economic divide? And isn't Google actually shaping the interests of its map readers, rather than showing them what is interesting?⁷⁸

reCAPTCHA

The World Wide Web is full of robots doing chores. My personal blog⁷⁹ for example, gets a few visits a day from Google's web crawler coming to check if there is anything new to index. Many of these robots have nefarious purposes. For instance, there are programs on the loose filling in web forms all over the internet to try and get their information listed for spam purposes or to find a weak spot in the technology and break into the server.⁸⁰ This is why often you have to prove that you are a human by doing a chore that is relatively easy for humans to do, while being difficult for robots. Typically reading a set of distorted letters and typing those in a form field. These challenges are named CAPTCHAs.⁸¹

In 2007, the computer scientist Luis von Ahn invented the reCAPTCHA as part of his work on human-based computation (in which machines outsource certain steps to humans). He thought it was a shame that the effort that people put into CAPTCHAs was wasted. In a reCAPTCHA people were shown two words out of old books that had been scanned by the Internet Archive: one word that reCAPTCHA already understood (to check if the person was indeed a human) and another that reCAPTCHA wasn't yet too sure about (to help digitize these books).⁸²

Google bought reCAPTCHA in 2009⁸³ and kept the service free to use for any website owner. They also switched the digitization effort from the open Internet Archive to its own proprietary book scanning effort. More than a million websites have currently integrated reCAPTCHA into their pages to check if their visitors are human. Google has a completely dominant market position for this service, as there are very few good alternatives. In 2014, Google changed reCAPTCHA's slogan from "Stop Spam, Read Books" to "Tough on Bots, Easy on Humans,"⁸⁴ and at the same time changed the problem from text recognition to

⁷⁸Bliss, "The Real Problem With 'Areas of Interest' on Google Maps."

⁷⁹De Zwart, "Medium Message – Writings by Hans de Zwart."

⁸⁰Often, to then use the server for mining cryptocurrencies.

⁸¹It stands for "Completely Automated Public Turing test to tell Computers and Humans Apart".

⁸²Thompson, "For Certain Tasks, the Cortex Still Beats the CPU."

⁸³Von Ahn and Cathcart, "Teaching Computers to Read."

⁸⁴"reCAPTCHA."

image recognition. In the current iteration, people have to look at a set of photos and click on all the images that have a traffic sign or have store fronts on them (see fig. 1 for an example).

With the switch to images, you no longer are helping Google to digitize books, you are now a trainer for its image recognition algorithms. As Google notes on its reCAPTCHA website under the heading “Creation of Value. Help everyone, everywhere – One CAPTCHA at a time.”:

Millions of CAPTCHAs are solved by people every day. reCAPTCHA makes positive use of this human effort by channeling the time spent solving CAPTCHAs into digitizing text, annotating images, and building machine learning datasets.

This in turn helps preserve books, improve maps, and solve hard AI problems.⁸⁵

Gabriela Rojas-Lozano has tried to sue Google for making her do free labor while signing up for Gmail without telling her that she was doing this labor.⁸⁶ She lost the case because the judge was convinced that she still would have registered for a Gmail account even if she had known about giving Google the ten seconds of labor.⁸⁷ Her individual “suffering” was indeed ludicrous, but she did have a point if you look at society at large. Every day, hundreds of millions of people fill in reCAPTCHAs for Google to prove that they are human.⁸⁸ This means that all of us give Google more than 135,000 FTE of our labor for free.⁸⁹ Google’s topnotch image recognition capability is partially enabled—and has certainly been catalysed—by this free labor.

In June of 2018, Gizmodo reported that Google had contracted with the United States Department of Defense to “help the agency develop artificial intelligence for analyzing drone footage.”⁹⁰ This led to quite an outrage among Google employees, who weren’t happy with their company offering surveillance resources to the military. I personally was quite upset from the idea that all my clicking on store fronts (as I am regularly forced to do, to access the information that I need, even on services that have nothing to do with Google), is now helping the US with its drone-based assassination programs in countries like Afghanistan, Yemen and Somalia.⁹¹

⁸⁵“reCAPTCHA - Creation of Value.”

⁸⁶Harris, “Massachusetts Woman’s Lawsuit Accuses Google of Using Free Labor to Transcribe Books, Newspapers.”

⁸⁷Dinzeo, “Google Ducks Gmail Captcha Class Action.”

⁸⁸“What Is reCAPTCHA?”

⁸⁹Assuming 1,500 working hours per year and 200 million reCAPTCHAs filled in per day, taking 10 seconds each. This estimate is likely to be too low, but probably is at the right order of magnitude.

⁹⁰Conger and Cameron, “Google Is Helping the Pentagon Build AI for Drones.”

⁹¹Scahill, “The Assassination Complex.”

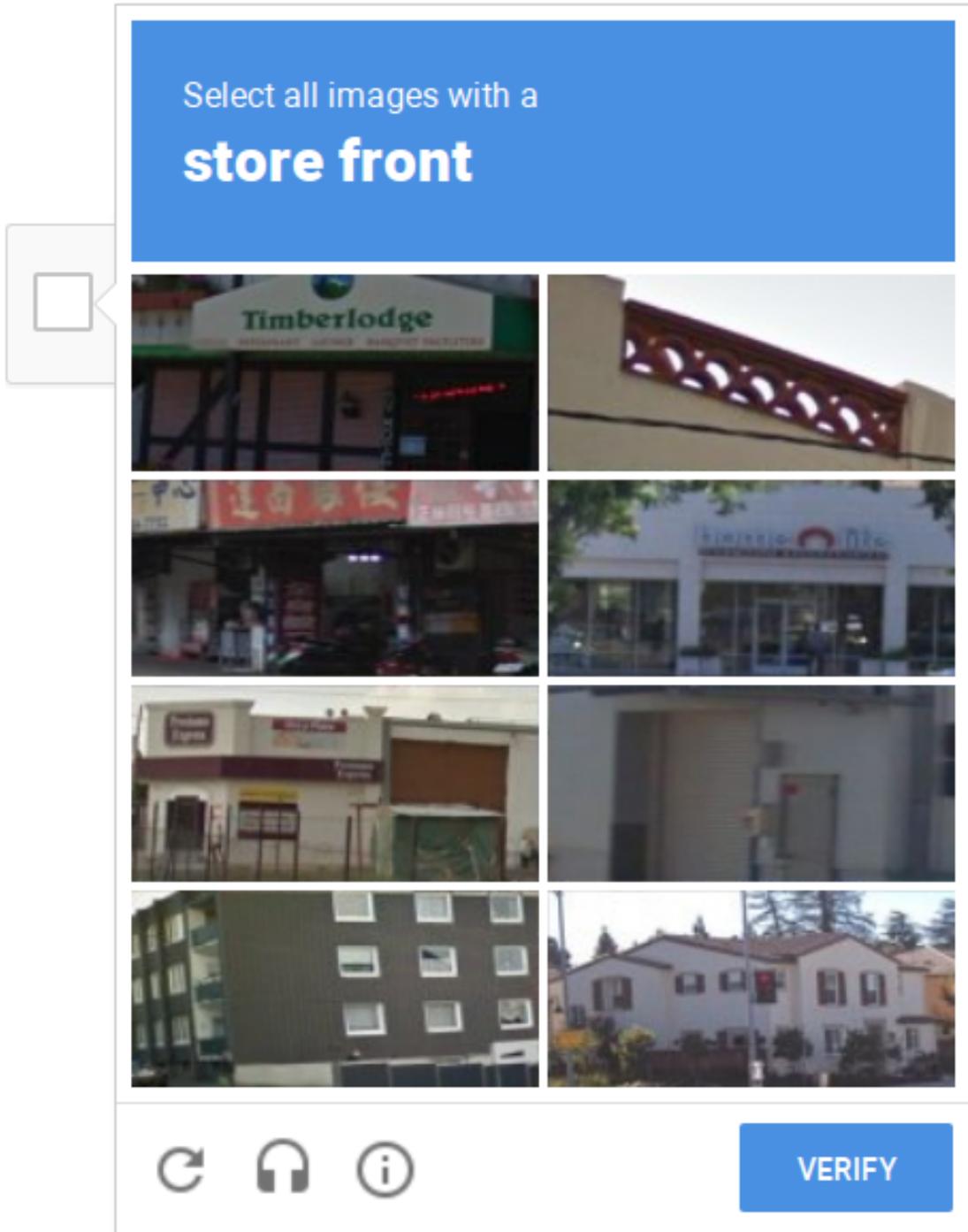


Figure 1: Google's reCAPTCHA asking to identify store fronts

Part 2: How is that problematic?

Now that we have a clear idea about our technological predicament, we can start to explore the potential effects that this might have on the structure of our society. It is obvious that these effects will be far-reaching, but at the same time they are undertheorized. As Zuboff writes:

We've entered virgin territory here. The assault on behavioral data is so sweeping that it can no longer be circumscribed by the concept of privacy and its contests. This is a different kind of challenge now, one that threatens the existential and political canon of the modern liberal order defined by principles of self-determination that have been centuries, even millennia, in the making. I am thinking of matters that include, but are not limited to, the sanctity of the individual and the ideals of social equality; the development of identity, autonomy, and moral reasoning; the integrity of contract, the freedom that accrues to the making and fulfilling of promises; norms and rules of collective agreement; the functions of market democracy; the political integrity of societies; and the future of democratic sovereignty.⁹²

I will look at the three features of our technological predicament through the lens of justice as fairness and freedom as non-domination. In both cases, I come to the conclusion that the effects are deleterious. Data-driven appropriation leads to injustices, whereas the domineering scale and the asymmetrical relationships negatively affect our freedom.

Injustice in our technological predicament

To assess whether our technological predicament is just, we will look at it from the perspective of Rawls's principles of justice. There are three central problems with the basic structure in our digitizing society. The first is a lack of equality in the division of the basic liberties, the second is an unjust division of both public and primary goods, and a final problem is tech's reliance on utilitarian ethics to justify their behavior.

The demands of justice as fairness

For John Rawls, the subject of justice is what he calls the "basic structure of society", which is "the way in which the major social institutions distribute fundamental rights and duties

⁹²Zuboff, "Google as a Fortune Teller."

and determine the division of advantages from social cooperation.”⁹³ Major social institutions are the principal economic and social arrangements and the political constitution.

The expository and intuitive device that Rawls uses to ensure that his conception of justice is fair is the “original position”. He writes: “One conception of justice is more reasonable than another, or justifiable with respect to it, if rational persons in the initial situation would choose its principles over those of the other for the role of justice.”⁹⁴ The restrictions that the original position imposes on the arguments for principles of justice help with the justification of this idea:

It seems reasonable and generally acceptable that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles. It also seems widely agreed that it should be impossible to tailor principles to the circumstances of one’s own case. We should insure further that particular inclinations and aspirations, and persons’ conceptions of their good do not affect the principles adopted. The aim is to rule out those principles that it would be rational to propose for acceptance [...] only if one knew certain things that are irrelevant from the standpoint of justice. [...] To represent the desired restrictions one imagines a situation in which everyone is deprived of this sort of information. One excludes the knowledge of those contingencies which sets men at odds and allows them to be guided by their prejudices. In this manner the veil of ignorance is arrived at in a natural way.”⁹⁵

The parties in the original position, and behind this veil of ignorance, are to be considered as equals. Rawls: “The purpose of these conditions is to represent equality between human beings as moral persons, as creatures having a conception of their good and capable of a sense of justice.”⁹⁶

According to Rawls, there would be two principles of justice that “rational persons concerned to advance their interests would consent to as equals when none are known to be advantaged or disadvantaged by social and natural contingencies.”⁹⁷ The first principle requires equality in the assignment of basic rights and duties:

Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all.⁹⁸

⁹³Rawls, *A Theory of Justice*, 6.

⁹⁴Ibid., 15–16.

⁹⁵Ibid., 16–17.

⁹⁶Ibid., 17.

⁹⁷Ibid., 17.

⁹⁸Ibid., 266.

Whereas the second principle holds that social and economic inequalities are only just if they result in compensating benefits for everyone and the least advantaged in particular:

Social and economic inequalities are to be arranged so that they are both:

- (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and
- (b) attached to offices and positions open to all under conditions of fair equality of opportunity.⁹⁹

These principles are to be ranked in lexical order. This means that the basic liberties can only be restricted for the sake of liberty (so when the less extensive liberty strengthens the total system of liberties shared by all, or when the less than equal liberty is acceptable to those with the lesser liberty), and that the second principle of justice goes before the principle of efficiency and before the principle of maximizing the sum of advantages.¹⁰⁰

For Rawls, the second principle expresses an idea of reciprocity. Even though the principle initially looks biased towards the least favored, Rawls argues that “the more advantaged, when they view the matter from a general perspective, recognize that the well-being of each depends on a scheme of social cooperation without which no one could have a satisfactory life; they recognize also that they can expect the willing cooperation of all only if the terms of the scheme are reasonable. So they regard themselves as already compensated [...] by the advantages to which no one [...] had a prior claim.”¹⁰¹

Lack of equality

To show how data-driven appropriation leads to inequality, I will use the investigative journalism of political science professor Virginia Eubanks. She has published her research in *Automating Inequality*.¹⁰² According to Eubanks:

Marginalized groups face higher levels of data collection when they access public benefits, walk through highly policed neighborhoods, enter the health-care system, or cross national borders. That data acts to reinforce their marginality when it is used to target them for suspicion and extra scrutiny. Those groups seen as undeserving are singled out for punitive public policy and more intense surveillance, and the cycle begins again. It is a kind of collective red-flagging,

⁹⁹Ibid., 266.

¹⁰⁰Ibid., 266.

¹⁰¹Ibid., 88.

¹⁰²Eubanks, *Automating Inequality*.

a feedback loop of injustice.¹⁰³

She argues that we have forged “a *digital poorhouse* from databases, algorithms, and risk models,”¹⁰⁴ and demonstrates this by writing about three different government programs that exhibit these features: a welfare reform effort, an algorithm to distribute subsidized houses to homeless people, and a family screening tool. The latter gives the most clear example of the possible unjust effects of recursively using data to create models.

The Allegheny Family Screening Tool (AFST) is an algorithm—based on machine learning—that aims to predict which families are at a higher risk of abusing or neglecting their children.¹⁰⁵ The Allegheny County Department of Human Services has created a large warehouse combining the data from twenty-nine different government programs, and has bought a predictive modelling methodology based on research in New Zealand¹⁰⁶ to use this data to make predictions of risk.

There is a lot of room for subjectivity when deciding what is to be considered neglect or abuse of children. “Is letting your children walk to a park down the block alone neglectful?”, Eubanks asks.¹⁰⁷ Where to draw the line between neglect and conditions of poverty is particularly difficult.¹⁰⁸ Eubanks is inspired by Cathy O’Neil, who says that “models are opinions embedded in mathematics,”¹⁰⁹ to do a close analysis of the AFST algorithm. She finds some serious design flaws that limit its accuracy:

It predicts referrals to the child abuse and neglect hotline and removal of children from their families—hypothetical proxies for child harm—not actual child maltreatment. The data set it utilizes contains only information about families who access public services, so it may be missing key factors that influence abuse and neglect. Finally, its accuracy is only average. It is guaranteed to produce thousands of false negatives and positives annually.¹¹⁰

The use of public services as an input variable means that low-income people are disproportionately represented in the database. This is because professional middle class families mostly rely on private sources for family support. Eubanks writes: “It is interesting to imagine the response if Allegheny County proposed including data from nannies, babysitters, private therapists, Alcoholics Anonymous, and luxury rehabilitation centers to predict

¹⁰³Ibid., 6–7.

¹⁰⁴Ibid., 12–13.

¹⁰⁵Ibid., 127–73.

¹⁰⁶Vaithianathan et al., “Children in the Public Benefit System at Risk of Maltreatment.”

¹⁰⁷Eubanks, *Automating Inequality*, 130.

¹⁰⁸Ibid., 130.

¹⁰⁹O’Neil, *Weapons of Math Destruction*, 21.

¹¹⁰Eubanks, *Automating Inequality*, 146.

child abuse among wealthier families.”¹¹¹ She calls the current program a form of “poverty profiling”:

Like racial profiling, poverty profiling targets individuals for extra scrutiny based not on their behavior but rather on a personal characteristic: living in poverty. Because the model confuses parenting while poor with poor parenting, the AFST views parents who reach out to public programs as risks to their children.¹¹²

Eubanks’s conclusion about automated decision-making on the basis of the three examples in her book is damning:

[It] shatters the social safety net, criminalizes the poor, intensifies discrimination, and compromises our deepest national values. It reframes shared social decisions about who we are and who we want to be as systems engineering problems. And while the most sweeping digital decision-making tools are tested in what could be called “low rights environments” where there are few expectations of political accountability and transparency, systems first designed for the poor will eventually be used on everyone.¹¹³

Eubanks’s examples all relate to how the state interferes with its citizens rights. These examples are still relevant to this thesis because they clearly show what happens when algorithms and data are used to make decisions about people and what these people are entitled to. The processes of the state at least have a level of accountability and the need for legitimacy in their decision making. The same can’t be said for accumulators like Google and Facebook. They are under no democratic governance and don’t have any requirements for transparency. This makes it harder to see the unequal consequences of their algorithmic decision making, and as a result makes it harder to question those.

One example of an unequal treatment of freedom of speech was highlighted by ProPublica in a investigative piece titled *Facebook’s Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children*.¹¹⁴ ProPublica used internal documents from Facebook to shed light on the algorithms that Facebook’s censors use to differentiate between hate speech and legitimate political expression.

The documents suggested that “at least in some instances, the company’s hate-speech rules tend to favor elites and governments over grassroots activists and racial minorities. In so

¹¹¹Ibid., 157.

¹¹²Ibid., 158.

¹¹³Ibid., 12.

¹¹⁴Angwin and Grassegger, “Facebook’s Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children.”

doing, they serve the business interests of the global company, which relies on national governments not to block its service to their citizens.”¹¹⁵ Up until very recently, Facebook did not publish the enforcement guidelines for its “community standards,”¹¹⁶ only after increased pressure from civil society did it decide to be transparent about its rules.¹¹⁷ There are endless examples of marginalized groups who have lost their audience because Facebook has decided to block their posts or their pages. Often they are the victims of a rule where Facebook protects certain special categories (like ethnicity or gender), but not subsets of those categories. This has led to the absurd situation where “white men” is a category that is protected from hate speech, but “black children” or “female drivers” are not.¹¹⁸

There is some proof that Facebook uses the profitability of a particular page or post as one of the criteria in the decision making about whether to remove it or not. When a Channel 4 reporter went undercover, he was told by his content moderation trainer that the page of the extreme right organization Britain First was left up—even though they had broken the rules many times—because “they have a lot of followers so they’re generating a lot of revenue for Facebook.”¹¹⁹ A Dutch Facebook moderator in Berlin had a similar story about the hate speech which was directed at the black Dutch politician and activist Sylvana Simons. He wasn’t allowed to remove any of it, mainly because Facebook has no incentive to take down content. This changed when the reporting about Simons turned on Facebook itself.¹²⁰ As soon as there is media attention for a particular decision, Facebook will often change course.

Kate Klonick, an academic specializing in corporate censorship, fears that Facebook is evolving into a place where celebrities, world leaders, and other prominent people “are disproportionately the people who have the power to update the rules.”¹²¹ This is a form of class justice and a clear example of a lack of equality. Dave Willner, a former member of Facebook’s content team, makes the explicit connection with justice. He says that Facebook’s approach is “more utilitarian than we are used to in our justice system, [...] it’s fundamentally not rights-oriented.”¹²²

¹¹⁵Ibid.

¹¹⁶Facebook’s euphemism for a code of conduct.

¹¹⁷Bickert, “Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process.”

¹¹⁸Angwin and Grassegger, “Facebook’s Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children.”

¹¹⁹Hern, “Facebook Protects Far-Right Activists Even After Rule Breaches.”

¹²⁰Kreling, Modderkolk, and Duin, “De Hel Achter de Façade van Facebook.”

¹²¹Angwin and Grassegger, “Facebook’s Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children.”

¹²²Ibid.

Abuse of the commons

For what I've been calling "appropriation", Rawls would most likely use the Aristotelian term "*pleonexia*", which he defines as "gaining some advantage for oneself by seizing what belongs to another, his¹²³ property, his reward, his office, and the like, or by denying a person that which is due to him, the fulfillment of a promise, the repayment of a debt, the showing of proper respect, and so on."¹²⁴ For Rawls, it is clear: "We are not to gain from the cooperative labors of others without doing our fair share."¹²⁵

Google using the collaborative effort of Wikipedia for their own gains (as seen in the YouTube case study above), is of course one of the more obvious examples of what Rawls does not allow. While researching this thesis, I encountered another dreadful example of this mechanism.¹²⁶

Google augments their search results for certain keywords with an information box, for which the information comes from Wikipedia. It does this to quickly provide the information most searchers will be looking for, and thus to make their search engine even more attractive. Google does not compensate Wikipedia for this use.¹²⁷ When I searched for "Los Angeles" in late July of 2018, it showed me an information box with despicable racist contents (see fig. 2 for a censored version of what I saw). After a bit of research, I found out that Wikipedia's page had been vandalized at 8:00 in the morning, and that a Wikipedia volunteer had cleaned up the mess one hour later. Google had indexed the vandalized page, but hadn't yet indexed the cleaned up version, even though it was ten hours after the problem had been fixed.

Because of Google's dominance, it is reasonable to assume that way more people will see this vandalized version of the information on the Google page than on the Wikipedia page. It is probable that the vandal's main purpose was to influence Google's search results. If that is indeed the case, then it means that Google using Wikipedia to spruce up their results has a detrimental effect on the quality of the collaborative encyclopedia. The fact that a Republican senator thought he had to publicly prove that he was still alive, after Google erroneously listed him as having passed away, proves that point. This mistake was also the result of a vandal making a change in Wikipedia, but this fact was nowhere mentioned in

¹²³Regrettably, there are only ever male protagonists in Rawls's examples.

¹²⁴Rawls, *A Theory of Justice*, 9.

¹²⁵Ibid., 96, 301.

¹²⁶I published the following story as an opinion piece in the NRC newspaper. See: De Zwart, "Miljar-denbedrijf Google Geeft Geen Cent Om de Waarheid."

¹²⁷To be clear: Wikipedia does allow the free use of their information under a Creative Commons Attribution-ShareAlike 3.0 license. Google complies with the attribution clause, but fails to tell its visitors under what license the information is available. See: "Text of Creative Commons Attribution-ShareAlike 3.0 Unported License."



Los Angeles
Stad in Californië

Kill all the filthy, evil ██████ and kill all ██████ and kill all ██████ there.
Kill all sh*t skinned ██████ and kill all ██████ [Wikipedia](#)

Weer: 24 °C, wind ZW met 5 km/h, 62% luchtvochtigheid
Lokale tijd: zaterdag 10:16
Bevolking: 3,976 miljoen (2016)

Een reis plannen

 Reisgids voor Los Angeles

 3-sterrenhotel gemiddeld € 193, 5-sterrenhotel gemiddeld € 628

Figure 2: Google.nl result for the search term “Los Angeles” on July 21st, 2018

the media coverage of the event.¹²⁸

This is a straightforward example of an accumulator appropriating the commons. There are two more complex (and more impactful) ways that our technological predicament is enabling the abuse of the commons. The first, is how the predictive knowledge about how the world works is being enclosed, a problem with the informational commons. The second, is the way our attention is taken away from us, a problem with the attentional commons.

The informational commons

In 1993, Bruce Sterling wrote beautifully about the internet as a public good, comparing the anarchical nature of the internet with the way the English language develops:

Nobody rents English, and nobody owns English. As an English-speaking person, it's up to you to learn how to speak English properly and make whatever use you please of it [...]. Otherwise, everybody just sort of pitches in, and somehow the thing evolves on its own, and somehow turns out workable. And interesting. Fascinating, even. [...] "English" as an institution is public property, a public good. Much the same goes for the Internet. [...] It's an institution that resists institutionalization. The Internet belongs to everyone and no one.¹²⁹

Rawls writes about public goods in the context of looking at economic systems to see if they can satisfy the two principles of justice. According to him, they have two characteristic features: indivisibility and publicness. Public goods "cannot be divided up as private goods can and purchased by individuals according to their preferences for more or less."¹³⁰ Rawls acknowledges the free-rider problem (individuals avoiding doing their share), and how this will limit the chances for voluntary agreements about the public good to develop. He also sees clearly how the externalities of the production of public goods will not be reckoned with by the market.¹³¹ So for him it is evident "that the indivisibility and publicness of certain essential goods, and the externalities and temptations to which they give rise, necessitate collective agreements organized and enforced by the state. [...] Some collective arrangement is necessary and everyone wants assurance that it will be adhered to if he is willingly to do his part."¹³²

Current literature sees public goods as one form of a commons, "a resource shared by a

¹²⁸See for example: Ehrlich, "GOP Senator Says He Is Alive Amid Google Searches Suggesting He Is Dead."

¹²⁹Sterling, "Science Column #5 'Internet'."

¹³⁰Rawls, *A Theory of Justice*, 235.

¹³¹Rawls: "There are the striking cases of public harms, as when industries sully and erode the natural environment." See: *ibid.*, 237

¹³²*Ibid.*, 237.

group of people that is subject to social dilemmas.”¹³³ Charlotte Hess and Elinor Ostrom use two dimensions to categorize the different forms of commons. The first is “exclusion”, how difficult or easy is it to stop somebody from accessing the resource (similar to Rawls’s publicness). And the second is “subtractability”, does the use by one person subtract from the available goods for others (this comes close to Rawls’s indivisibility, but is more often framed as rivalry). Public goods are those with a low subtractability and a high difficulty for exclusion.¹³⁴

One of the issues with commons is always the threat of enclosure. As Hess and Ostrom write: “The narrative of the enclosure is one of privatization, the haves versus the have-nots, the elite versus the masses.”¹³⁵ The first enclosure¹³⁶ was the withdrawing of community rights (by landowners and the state) from the European shared agricultural fields.¹³⁷

In 2007, James Boyle argued that we were in the second enclosure movement, which he grandiloquently¹³⁸ called “the enclosure of the intangible commons of the mind.”¹³⁹ Hess and Ostrom think that “this trend of enclosure is based on the ability of new technologies to ‘capture’ resources that were previously unowned, unmanaged, and thus, unprotected.”¹⁴⁰ Boyle focuses on the use of intellectual property rights and the related enforcement technologies to clamp down on the ease of copying. Basically the idea of enclosure is always to make it easier to exclude people from accessing the resource. Public goods that have become easy to exclude turn into toll or club goods.¹⁴¹

I want to argue that the data-driven appropriation in our technological predicament is a third movement of enclosure, turning public goods into toll goods. Going forward, it isn’t just intellectual labor (the “commons of the mind”) that is enclosed. It is our actual understanding of the world—an understanding that is predicated on measuring our social, cultural and economic behaviors—for which access becomes exclusive and under the terms of the data accumulators. This is happening in domain after domain; whether it is transportation, health, education or communication.

It is hard to find a precise enough analogy to make it easier to understand how this third movement of enclosure works. But maybe a look at how we predict the weather can help. Currently, gathering the weather data and using this to turn it into (predictive) models is

¹³³Hess and Ostrom, “An Overview of the Knowledge Commons,” 3.

¹³⁴Ibid., 8–9.

¹³⁵Ibid., 12.

¹³⁶The authors probably mean that this was the first enclosure movement to be theorized.

¹³⁷Ibid., 12.

¹³⁸He himself grandiloquently used the word “grandiloquently”.

¹³⁹Boyle, “The Second Enclosure Movement,” 19.

¹⁴⁰Hess and Ostrom, “An Overview of the Knowledge Commons,” 12.

¹⁴¹Ibid., 9.

mostly a public and decidedly collective effort. There are some private companies that help with collecting the data (airline companies for example), and the data is mostly freely available for anybody to use for their own purposes. If we would apply the third enclosure model to this situation, then an accumulator would come in and would outsource all the measuring and data collection to private individuals and small businesses (sometimes without them even knowing, and occasionally in return for access to some information). The accumulator would then use this data to create predictive weather models and would share parts of these predictions (basically when it suits their purposes, for example in order to get more sensor data) with the people who agree to their terms of service, or they would sell their predictions to public institutions as steering mechanisms for public policy. Some of these accumulators might even slightly adjust the predictions they share, in order to shape the behavior of the user of the prediction.

I therefore heartily agree with Aral Balkan’s forceful critique of swapping out a public goods based infrastructure for a toll based one, and I imagine Rawls would agree with him too:

It is not the job of a corporation to “develop the social infrastructure for community” as Mark [Zuckerberg] wants to do. Social infrastructure must belong to the commons, not to giant monopolistic corporations like Facebook. The reason we find ourselves in this mess with ubiquitous surveillance, filter bubbles, and fake news (propaganda) is precisely due to the utter and complete destruction of the public sphere by an oligopoly of private infrastructure that poses as public space.¹⁴²

The attentional commons

In his 2015 book *The World Beyond Your Head*,¹⁴³ Matthew Crawford makes a compelling case for an attentional commons. He considers our attention as a resource, because each of us only has so much of it. But it is a resource for which we currently lack what he calls a “political economy.”¹⁴⁴ Crawford explains how we hold certain resources—like the air we breathe and the water we drink—in common. We don’t pay a lot of attention to them (usually just taking them for granted), but it is their availability that makes everything else we do possible. Crawford thinks that “the absence of noise is a resource of just this sort. More precisely, the valuable thing that we take for granted is the condition of *not being addressed*. Just as clean air makes respiration possible, silence, in this broader sense, is what makes it possible to *think*.”¹⁴⁵

¹⁴²Balkan, “Encouraging Individual Sovereignty and a Healthy Commons.”

¹⁴³Crawford, *The World Beyond Your Head*.

¹⁴⁴Ibid., 11.

¹⁴⁵Ibid., 11.

It is clear that resources like water and air need robust regulations to be protected as common resources. In the absence of these regulations, they “will be used by some in ways that make them unusable for others—not because they are malicious or careless, but because they can make money using them this way. When this occurs, it is best understood as a *transfer of wealth* from ‘the commons’ to private parties.”¹⁴⁶

We have already reached the point where (cognitive) silence is offered as a luxury good. Basically, our attention is taken from us, and we then get to buy it back. Crawford gives the example of an airport, where he encounters ads inside his security tray and on the luggage belt,¹⁴⁷ but is completely liberated from this noise as soon as he steps into the business class lounge. Silence as a luxury is already part of our technological predicament too. YouTube offers a premium subscription for which the main benefit is that it is completely ad-free¹⁴⁸ and the Amazon Kindle e-readers come with “special offers”—Amazon’s euphemism for advertising—unless a one-time fee has been paid to remove them.¹⁴⁹

The era of accumulation makes the creating of a political economy for attention more urgent. We are increasingly the object of targeted attention grabbing. Crawford wants to supplement the right to privacy with a right not to be addressed: “This would apply not, of course, to those who address me face-to-face as individuals, but to those who never show their face, and treat my mind as a resource to be harvested by mechanized means.”¹⁵⁰

Crawford makes a beautiful argument why attention is both highly personal and intimate, while also being constitutive of our shared world:

Attention is the thing that is most one’s own: in the normal course of things, we choose what to pay attention to, and in a very real sense this determines what is real for us; what is actually present to our consciousness. Appropriations of our attention are then an especially intimate matter.

But it is also true that our attention is directed to a world that is shared; one’s attention is not simply one’s own, for the simple reason that its objects are often present to others as well. And indeed there is a moral imperative to pay attention to the shared world, and not get locked up in your own head. Iris Murdoch writes that to be good, a person “must know certain things about his surroundings, most obviously the existence of other people and their claims.”¹⁵¹

¹⁴⁶Ibid., 12.

¹⁴⁷Airports are one of the places with the most ads, mainly because it is “a high dwell time environment, delivering a captive audience.” See: Ruane, “Signs of Our Times.”

¹⁴⁸This doesn’t mean that you will no longer be tracked though. See: “YouTube Premium.”

¹⁴⁹Falcone, “Amazon Backtracks, Will Offer \$15 Opt-Out for Ads on Kindle Fire Tablets.”

¹⁵⁰Crawford, *The World Beyond Your Head*, 13.

¹⁵¹Ibid., 13–14.

This matches the two reasons he gives for finding the concept of a commons suitable in the context of a discussion about attention:

First, the penetration of our consciousness by interested parties proceeds very often by the *appropriation* of attention in public spaces, and second, because we rightly *owe* to one another a certain level of attentiveness and ethical care. The words italicized in the previous sentence rightly put us in a political economy frame of mind, if by “political economy” we can denote a concern for justice in the public exchange of some private resource.¹⁵²

I want to make the argument that it is beneficial to see attention analogously to a “primary good” in the Rawlsian sense of the word. Rawls uses the conception of “primary goods” as a way to address the practical political problem of people having conflicting comprehensive conceptions of the good. However distinct those conceptions may be, they require the same primary goods for their advancement “that is, the same basic rights, liberties, and opportunities, as well as the same all-purpose means such as income and wealth, all of which are secured by the same social bases of self-respect. These goods [...] are things that citizens need as free and equal persons, and claims to these goods are counted as appropriate claims.”¹⁵³ Rawls provides a basic lists of primary goods under five headings:

(i) basic rights and liberties, of which a list may also be given; (ii) freedom of movement and free choice of occupation against a background of diverse opportunities; (iii) powers and prerogatives of offices and positions of responsibility in the political and economic institutions of the basic structure; (iv) income and wealth; and finally, (v) the social bases of self-respect.¹⁵⁴

Rawls allowed for things to be added to the list (if needed), for example to include other goods or maybe even to include mental states.¹⁵⁵ Attention is an all-purpose mean that citizens need to live out their conception of the good life. And the way that technology companies manage to appropriate this attention is leading to an unjust division of the resource.

The Center for Humane Technology (which used to be called Time Well Spent), lays out the problem with great clarity:

Facebook, Twitter, Instagram, and Google have produced amazing products that have benefited the world enormously. But these companies are also caught in a *zero-sum race for our finite attention*, which they need to make money.

¹⁵²Ibid., 14.

¹⁵³Rawls, “The Priority of Right and Ideas of the Good,” 257.

¹⁵⁴Ibid., 257.

¹⁵⁵Rawls himself mentions leisure time and the absence of physical pain as potential candidates. See: *ibid.*, 257

Constantly forced to outperform their competitors, they must use increasingly persuasive techniques to keep us glued. They point AI-driven news feeds, content, and notifications at our minds, continually learning how to hook us more deeply—from our own behavior. [...] These are not neutral products. *They are part of a system designed to addict us.*¹⁵⁶

A utilitarian ethics

Data-driven accumulators are starting to realize that they need to justify the disparate impact that their data-driven technologies have. Google for example, acknowledges the power of artificial intelligence as a technology, and understands that the technology will have a significant impact on our society. To address their “deep responsibility”, they published a set of seven principles¹⁵⁷ that guide their artificial intelligence work.¹⁵⁸

The principles make it clear what type of ethical stance lies behind Google’s approach. The first principle is called “Be socially beneficial” and reads as follows:

The expanded reach of new technologies increasingly touches society as a whole. Advances in AI will have transformative impacts in a wide range of fields, including healthcare, security, energy, transportation, manufacturing, and entertainment. As we consider potential development and uses of AI technologies, we will take into account a broad range of social and economic factors, and *will proceed where we believe that the overall likely benefits substantially exceed the foreseeable risks and downsides.*¹⁵⁹

This is clearly a utilitarian perspective: they will use artificial intelligence as long as the benefits exceed the risks. Arguably it would be hard for them to espouse anything but a teleological ethical theory. They need it to justify what they are doing. Unfortunately, they don’t describe their utility function. What is to be considered socially beneficial, and what is seen as a cost to society? How will this utility be quantified, and who gets to decide on these questions?

Rawls understands the intuitive appeal of a utilitarian ethics. In a teleological approach you can define the good independently from the right, and then define the right as that which maximizes the good. The appeal to Google is clear, not only because they couldn’t justify their behaviour with a deontological approach, but also because utilitarianism seemingly

¹⁵⁶“Our Society Is Being Hijacked by Technology.”

¹⁵⁷Pichai, “AI at Google.”

¹⁵⁸The principles were probably a direct reaction to their employees protesting a Google contract with the US Department of Defense. See: Shane and Wakabayashi, “‘The Business of War’.”

¹⁵⁹Pichai, “AI at Google,” italics added.

embodies rationality. “It is natural to think that rationality is maximizing something and that in morals it must be maximizing the good. Indeed, it is tempting to suppose that it is self-evident that things should be arranged so as to lead to the most good.”¹⁶⁰

For Rawls, “the striking feature of the utilitarian view of justice is that it does not matter [...] how this sum of satisfactions is distributed among individuals [...]. The correct distribution [...] is that which yields the maximum fulfillment.”¹⁶¹ He therefore dismisses the principle of utility as “inconsistent with the idea of reciprocity implicit in the notion of a well-ordered society.”¹⁶² According to Rawls, the utilitarian view of “social cooperation is the consequence of extending to society the principle of choice for one man, and then, to make this extension work, conflating all persons into one through the imaginative acts of the impartial sympathetic spectator.”¹⁶³ He sees no reason why from the original position of equality this option would be seen as acceptable: “Since each desires to protect his interests, his capacity to advance the conception of the good, no one has a reason to acquiesce in an enduring loss for himself in order to bring about a greater net balance of satisfaction.”¹⁶⁴ Rawls finishes his treatment of classic utilitarianism with a damning indictment:

Utilitarianism does not take seriously the distinction between persons.¹⁶⁵

Rawls’s description of utilitarianism matches one for one with the espoused ethical theory of most technology companies. I therefore think the following is true and helps to elucidate our technological predicament:

Google does not take seriously the distinction between persons.¹⁶⁶

Unfreedom in our technological predicament

Now that it is clear that the current ecosystems of data appropriation have many unjust consequences, I want to argue a less obvious point. Even though many of the technologies, enabled by data, create new options for us and increase our choices, they actually make us less free. To maintain that position, I will use a particular conception of freedom: civic republicanism.¹⁶⁷ As this idea of freedom is most eloquently explained by Philip Pettit, I will stay very close to his reasoning.

¹⁶⁰ Rawls, *A Theory of Justice*, 22.

¹⁶¹ *Ibid.*, 23.

¹⁶² *Ibid.*, 13.

¹⁶³ *Ibid.*, 24.

¹⁶⁴ *Ibid.*, 13.

¹⁶⁵ *Ibid.*, 24.

¹⁶⁶ Google is just the example here, because they’ve made their ethics explicit. Many of the other technology companies behave on the basis of a similar ethical stance.

¹⁶⁷ Also called “neorepublicanism”.

The demands of freedom as non-domination

To illustrate the crucial point about his conception of freedom, Pettit often uses *A Doll's House* as an example. The protagonists in this classic Ibsen play are Torvalds, a young banker, and his wife, Nora. During the late 19th century a husband had near limitless power over his wife, but Torvalds completely dotes over Nora, and denies her absolutely nothing. In practical daily life, she can basically do what she wants. According to Pettit, Nora might have many benefits, but you can't say she enjoys freedom in her relationship with Torvalds:

His hands-off treatment means that he does not interfere with her, as political philosophers say. He does not put any prohibitions or penalties in the way of her choices, nor does he manipulate or deceive her in her exercise of those choices. But is this enough to allow us to think of Nora as a free agent? If freedom consists in noninterference, as many philosophers hold, we must say that it is. But I suspect that like me, you will balk at this judgment. You will think that Nora lives under Torvald's thumb. She is the doll in a doll's house, not a free woman.¹⁶⁸

This becomes abundantly clear in the last act of the play, where Torvalds first forgives Nora for the sins she has committed (all with the purpose of helping to solve the problems of his making), and then tells her:

There is something so indescribably sweet and satisfying, to a man, in the knowledge that he has forgiven his wife—forgiven her freely, and with all his heart. It seems as if that had made her, as it were, doubly his own; he has given her a new life, so to speak; and she is in a way become both wife and child to him. So you shall be for me after this, my little scared, helpless darling.¹⁶⁹

That is when Nora realizes that Torvalds truly is a stranger to her. Soon after, she decides to leave him and the children behind. Her final action in the play is to slam the door as she leaves.

Pettit uses this example to show that the absence of interference isn't enough to make us free.¹⁷⁰ You also need “the absence of domination: that is, the absence of subjection to the will of others [...]”¹⁷¹ Pettit argues that your freedom should have depth (freedom as a property of choices) and that you must have this deep freedom over a broad range of choices

¹⁶⁸Pettit, *Just Freedom* xiv.

¹⁶⁹Ibsen, *A Doll's House*, 88.

¹⁷⁰In my Bachelor's thesis, I've attempted to show that a classic liberal negative conception of freedom as non-interference has a much harder time showing what's wrong with our current technological predicament than the republican ideal of freedom. See: De Zwart, “Liberty, Technology and Democracy.”

¹⁷¹Pettit, *Just Freedom* xv.

(freedom as a property of persons).

Freedom with depth

When is your freedom “deep”? If you have different options, what are then the conditions ensuring that your choice between those options is a free choice? Pettit has three conditions:

You enjoy freedom of choice between certain options to the extent that:

1. you have the room and the resources to enact the option you prefer,
2. whatever your own preference over those options, and
3. whatever the preference of any other as to how you should choose.¹⁷²

Having the room to enact the option you prefer means that there should be no interference with your options. That interference can be done in multiple ways. The option may be removed (blocking the ability to make the choice), the option may be replaced (by penalizing or burdening it), or the option may be misrepresented (deception about the available alternatives or manipulating the perception of the alternatives). There are certain ways in which a choice can be influenced without there being interference of this type. Incentivizing, persuading and nudging (without deception) all do not constitute interference because they do not remove, replace or misrepresent an option.¹⁷³

Next to having the room, you should also have the resources. If you lack all the necessary resources to be able to choose a certain option, then you can't be free to choose that option. Pettit categorizes resources into three broad areas: personal (the mental and bodily ability and knowhow needed to make the choice), natural (the conditions in the environment that put the option within reach) and social (the conventions and shared awareness that makes acts of communication possible).¹⁷⁴

The second clause says that you only enjoy freedom in your choice of options, if all the options are available in the ways the first clause stipulates, regardless of your own preference for any of the options. Thomas Hobbes saw that differently. He thought that somebody is a free agent as long as that person “is not hindred to doe what he has a will to do.”¹⁷⁵ This idea leads to the absurd situation that you would be able to liberate yourself by adapting your preferences. Isaiah Berlin dismissed that idea beautifully:

To teach a man that, if he cannot get what he wants, he must learn to want only what he can get, may contribute to his happiness or his security; but it will not

¹⁷²Ibid., 30.

¹⁷³Ibid., 34–35.

¹⁷⁴Ibid., 36–38.

¹⁷⁵Hobbes, *Leviathan*, 146.

increase his civil or political freedom.¹⁷⁶

Pettit summarizes this second clause: “To have a free choice between certain options, you must be positioned to get whichever option you might want however unlikely it is that you might want it.”¹⁷⁷

It is the third clause that sets civic republicans apart. Pettit: “Your capacity to enact the option you prefer must remain in place not only if you change your mind about what to choose, but also if others change their minds as to what you should choose.”¹⁷⁸ The basic idea is that “you cannot be free in making a choice if you make it in subjection to the will of another agent, *whether or not you are conscious of the objection*.”¹⁷⁹ As Pettit writes:

The republican insight is that you will also be subject to my will in the case where I let you choose the option you prefer—and would have let you choose any option you preferred—but only because I happen to want you to enjoy such latitude.¹⁸⁰

If the third clause wouldn't be deemed necessary for freedom, then it would be possible to liberate yourself through ingratiating, which is as absurd as liberating yourself through preference adaptation. From the republican perspective, liberty means that you live on your own terms and are exempt from the dominion of another. That is even more important than having the required resource to enact on your preference. This means “that it is inherently worse to be controlled by the free will of another than to be constrained by a contingent absence of resources.”¹⁸¹ Pettit cites Kant as somebody who gives this idea prominence:

Find himself in what condition he will, the human being is dependent upon many external things. [...] But what is harder and more unnatural than this yoke of necessity is the subjection of one human being under the will of another. No misfortune can be more terrifying to one who is accustomed to freedom, who has enjoyed the good of freedom, than to see himself delivered to a creature of his own kind who can compel him to do what he will [...].¹⁸²

Freedom with breadth

If our freedom of choice is protected from domination, what should then be the range of

¹⁷⁶Berlin, “Two Concepts of Liberty,” 32.

¹⁷⁷Pettit, *Just Freedom*, 41.

¹⁷⁸Ibid., 41.

¹⁷⁹Ibid., 43, emphasis added.

¹⁸⁰Ibid., 43.

¹⁸¹For Pettit this explains why we feel resentment when we are controlled by another's will, and only exasperation when the constraints don't have anything to do with the will. See: *ibid.*, 215n28.

¹⁸²Kant, *Notes and Fragments*, 11.

decisions in which a freedom of this type should be available? According to Pettit, to be a free person in the republican conception requires you to be objectively secured against the intrusions of others, and subjectively that this security is a matter of common awareness: your status as a free person “must be salient and manifest to all.”¹⁸³ This is because the recognition of the protection of your rights, reinforces that protection. This status can only be available under “a public rule of law in which all are treated as equals.”¹⁸⁴ As Pettit sums up:

The republican ideal of the free citizen holds that in order to be a free citizen you must enjoy non-domination in such a range of choice, and on the basis of such public resourcing and protection, that you stand on a par with others. You must enjoy a freedom secured by public laws and norms in the range of the fundamental or basic liberties. And in that sense, you must count as equal with the best.¹⁸⁵

Pettit then derives the basic liberties that should be associated with a free civic status from this idea of what a free citizen should be: “The ceiling constraint is that the basic liberties should not include choices that put people at loggerheads with one another and force them into competition”, and the “floor constraint is that the basic liberties should encompass all the choices that are co-enjoyable in this sense, not just a subset of them.”¹⁸⁶

To be co-enjoyable by all, a choice must meet two conditions. The first, is that the choice must be co-exercisable in the sense that “people must be able to exercise any one of the choices in the set, no matter how many others are exercising it at the same time”. Secondly, the choice must be co-satisfying in the sense that “people must be able [...] to derive satisfaction from the exercise of any choice, no matter how many others are exercising that choice, or any other choice in the set [...]”¹⁸⁷

Pettit then argues that co-exercisable are basically those choices that you can do on your own, and that co-satisfactory choices exclude those that do harm to others, that lead to overpowering or destructive effects, and those where exercising the choice together is counterproductive. Which basic liberties will satisfy these requirements and constraints will differ with the cultural, technological and economical characteristics of a particular society. For Pettit, a society that provides this robust form of freedom will count as just, democratic and sovereign:

¹⁸³Pettit, *Just Freedom*, 57.

¹⁸⁴Ibid., 58.

¹⁸⁵Ibid., 60.

¹⁸⁶Ibid., 62.

¹⁸⁷Ibid., 62–63.

If the society entrenches each against the danger of interference from others in the domain of the basic liberties, then it will count plausibly as a just society. If this entrenchment is secured under a suitable form of control by the citizenry, then the society will count as properly democratic [...]. And if the international relations among peoples guard each against the danger of domination by other states or by non-state actors, then each people will have the sovereign freedom to pursue such justice and democracy [...].¹⁸⁸

The power to manipulate

Technological mediation virtualises the relationships between us and the rest of the world. In the most general terms, you could say that you interact with a third party who shapes and forms this virtual reality through which we connect and interact with other people and other objects. Schematically:

you <-> third party creating a virtual interface <-> something or somebody

It is a much easier to shape a virtual information-based reality than it is to shape a material atom-based reality (for Facebook to change the color of their website from blue to green requires a change in one line of code, whereas for Facebook to change the color of its offices from blue to green will take many days of work). Moreover, this reality can be shaped at the personal level (it is much easier for Facebook to personalize their website and show it to me in my favorite color, than it is for them to show me their offices in my favorite color).

When more of what we pay attention to in the world is technologically mediated by virtual third parties, we become more vulnerable to manipulation. These third parties have the ability to shape and form our personal reality in such a way that it serves their aims.

Natasha Dow Schüll has given a brilliantly telling example of this phenomenon in her book *Addiction by Design*, in which she does an anthropological exploration of the world of gambling machines in Las Vegas. She explains, how much easier it became for vendors of slot machines to get their players into the zone, once the faces of these machines became virtual instead of the physical reels that were used before:

Virtual reel mapping has been used not only to distort players' perception of games' *odds* but also to distort their perception of *losses*, by creating "near miss" effects. Through a technique known as "clustering," game designers map a disproportionate number of virtual reel stops to blanks directly adjacent to winning symbols on the physical reels so that when these blanks show up on

¹⁸⁸Ibid., 73.

the central payline, winning symbols appear above and below them far more often than by chance alone.¹⁸⁹

Technology companies do similar things to manipulate the behavior of their users. In their paper about digital market manipulation in the “sharing” economy, Ryan Calo and Alex Rosenblat lay out some evidence of manipulation from Uber. Riders for example are shown fake cars:

A user may open her app and see many vehicles around her, suggesting that an Uber driver is close by should she decide to hail one. [...] [However] the representation of nearby Uber cars can be illusory. Clicking the button to request an Uber prompts a connection to the nearest driver, who may be much further away. The consumer may then face a wait time as an actual Uber driver wends their way toward the pick up location. Those icons that appeared where cars were not present are familiar to some participants as “phantom cars.”¹⁹⁰

Virtual manipulations like these are hard to check and validate. How can an average user know whether a car that is shown in their phone in an app is actually there, or whether it is a virtual fake trying to lure them into ordering an taxi. Drivers (or as Uber calls them: customers) are manipulated too. For example, by Uber hiding information about the market place. Heat maps about where surge prices are have been made less accurate, and now function as “a behavioral engagement tool but can effectively operate as a bait-and-switch mechanism similar to the use of phantom cars to entice ride-hailers.”¹⁹¹ Calo and Rosenblat directly address what this means for freedom:

These constraints on drivers’ freedom to make fully informed and independent choices reflect the broad information and power asymmetries that characterize the relationship between Uber and its drivers and illustrate how the Uber platform narrows the choices that drivers are free to make.¹⁹²

Manipulation can also very effectively be done through adjusting the ranking of search results. Robert Epstein and Ronald E. Robertson have researched whether the ranking of search results could alter the preferences of undecided voters in democratic elections. They found that biased search rankings can shift the preferences of undecided voters by 20% or more, and that this bias can be masked, so that people aren’t aware of the manipulation. They conclude: “Given that many elections are won by small margins, our results suggest that a search engine company has the power to influence the results of a substantial number

¹⁸⁹Schüll, *Addiction by Design*, 92.

¹⁹⁰Calo and Rosenblat, “The Taking Economy,” 1655.

¹⁹¹Ibid., 1662.

¹⁹²Ibid., 1662.

of elections with impunity. The impact of such manipulations would be especially large in countries dominated by a single search engine company.”¹⁹³ They call this type of influence the search engine manipulation effect.

In an article for Politico, Epstein goes a step further and outlines what he considers to be three credible scenarios for how Google could decide a US presidential election. Google could make the executive decision to do this, there could be a rogue employee or group of employees who could implement a change in the algorithm, and finally, there be could a digital bandwagon effect where higher search activity creates higher search rankings, boosting voter interest, leading to higher search activity, and so on.¹⁹⁴

Google’s reply to Epstein was telling. They called Epstein’s work “a flawed elections conspiracy theory” and argued that they have “never ever re-ranked search results on any topic (including elections) to manipulate user sentiment.”¹⁹⁵ Although I have my doubts about the veracity of that statement, I also think it fails to address the core of Epstein’s worries. The question of whether Google will ever change the search results in order to get an election outcome that would suit Google’s purposes is a different question than whether Google has the power to do so, if they wanted to. From Pettit’s republican perspective, it is not relevant whether the domineering power is ever exercised in order to assess the extent of our freedom.

Epstein’s conclusion about our technological predicament is as follows:

We are living in a world in which a handful of high-tech companies, sometimes working hand-in-hand with governments, are not only monitoring much of our activity, but are also invisibly controlling more and more of what we think, feel, do and say. The technology that now surrounds us is not just a harmless toy; it has also made possible undetectable and untraceable manipulations of entire populations – manipulations that have no precedent in human history and that are currently well beyond the scope of existing regulations and laws.¹⁹⁶

In this time where we increasingly become dependent on virtual representations of our world, the ability to manipulate people in order to create the future that you want is a logical consequence of the ability to predict the future. To be able to *create* the future you need two things:

1. An understanding of the world, in the sense that you know which circumstances lead

¹⁹³Epstein and Robertson, “The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections” E4512.

¹⁹⁴Epstein, “How Google Could Rig the 2016 Election.”

¹⁹⁵Singhal, “A Flawed Elections Conspiracy Theory.”

¹⁹⁶Epstein, “How Google Could Rig the 2016 Election.”

to what types of behavior.

2. The ability to change the circumstances, so that you can bring about the circumstances that will lead to the behavior that you want to create.

The current best known example of a company which tries to exploit this mechanism, is Cambridge Analytica. They have argued that they were influential in getting people to vote for Trump and for the Brexit.¹⁹⁷ Even now that Cambridge Analytica has gone bankrupt, their websites are still full of phrases like “Cambridge Analytica uses data to change audience behavior”¹⁹⁸ and “We find your voters and move them to action. [...] By knowing your electorate better, you can achieve greater influence [...]”¹⁹⁹ In their Trump case study they write “Analyzing millions of data points, we consistently identified the most persuadable voters and the issues they cared about. We then sent targeted messages to them at key times in order to move them to action.”²⁰⁰ Here they clearly spell out the steps of accumulation, which can of course also be applied to other domains than commercial marketing or political campaigns.²⁰¹

Dependence on philanthropy

Most of the accumulators use their tremendous power and influence for philanthropic and social goals. Sometimes this is done very explicitly and without clear business goals, like Google.org’s investment of 1 billion U.S. dollars over a period of five years to improve education, economic opportunity, and inclusion.²⁰² Sometimes the social goals nicely align with the business goals, like with Facebook’s family of Internet.org projects with the mission to bring “internet access and the benefits of connectivity to the portion of the world that doesn’t have them.”²⁰³²⁰⁴ But mostly, these companies consider themselves to already have a positive influence on the world. They charge their customers (mostly businesses that want to advertise with them) and provide the services to their users for free.²⁰⁵

This is why it can be Google’s mission to “Organize the world’s information and make it

¹⁹⁷Cadwalladr and Graham-Harrison, “Revealed.”

¹⁹⁸“Data Drives All That We Do.”

¹⁹⁹“Data-Driven Campaigns.”

²⁰⁰“Donald J. Trump for President.”

²⁰¹It is important to realize that there is probably a big gap between the commercial sales language that Cambridge Analytica uses and the actual abilities of its products (see for example: Wong, “Cambridge Analytica-Linked Academic Spurns Idea Facebook Swayed Election.”). However, it is still early days in using data to drive behavior and the predictions don’t have to be perfect for there to be results that have an impact.

²⁰²“Our \$1 Billion Commitment to Create More Opportunity for Everyone.”

²⁰³“Our Mission.”

²⁰⁴Unfortunately, Internet.org seems to equate “internet access” to “access to Facebook”. See: Kreiken, “Humanitair-Vrijheids-Vrede-Mensenrechten-Project-Facebook.”

²⁰⁵Only in a monetary sense, of course.

universally accessible and useful” and to do this “Not just for some. For everyone.”²⁰⁶ And why Facebook’s Mark Zuckerberg, when asked why he doesn’t use Facebook to push social agenda issues, answers as follows:

I think the core operation of what you do should be aimed at making the change that you want. [...] What we are doing in making the world more open and connected, and now hopefully building some of the social infrastructure for a global community—I view that as the mission of Facebook.²⁰⁷

But depending on private philanthropy is very problematic from the perspective of (republican) freedom. It clientelizes the user and turns them into dependents. As Pettit writes:

If people depend in an enduring way on the philanthropy of benefactors, then they will suffer a clear form of domination. Their expectations about the resources available will shift, and this shift will give benefactors an effective power of interference in their lives.²⁰⁸

Arbitrary control

An simple, yet incredibly clear, example of the arbitrary nature of our relationship with the technology giants, can often be found in their terms of service. Before we do a close reading of Google’s terms over service,²⁰⁹ it is important to realize that these terms also apply to people’s Gmail accounts or their photos in Google Photos. Some would argue that the data that these services contain about you, actually is you:

Today, we are all cyborgs. This is not to say that we implant ourselves with technology but that we extend our biological capabilities using technology. We are sharded beings; with parts of our selves spread across and augmented by our everyday things.²¹⁰

So when we look at these terms, we need to realize that we are talking about services that are part of people’s identities.

Firstly, Google wants you to understand that they can stop providing the service to you at any time:

We may suspend or stop providing our Services to you if you do not comply with our terms or policies or if we are investigating suspected misconduct.

²⁰⁶“Our Company.”

²⁰⁷Safian, “Mark Zuckerberg On Fake News, Free Speech, And What Drives Facebook.”

²⁰⁸Pettit, *Just Freedom*, 88.

²⁰⁹All of the quotations of terms come from: “Google Terms of Service.”

²¹⁰Balkan, “The Nature of the Self in the Digital Age.”

It is hard to always be compliant with their terms or policies, because they reserve the right to change these without proactively noticing the user (as we shall see a bit later).

Next, they want to make sure that they carry no responsibility for what you do (or anybody else does for that matter) with your Google account:

You are responsible for the activity that happens on or through your Google Account.

This responsibility isn't shared with Google. So even if you get hacked without it being your fault,²¹¹ you are still liable for the damage that is done with your account.

Even though they leave the ownership of what gets uploaded to their services with you, they do make you give them a worldwide and everlasting license on your content. Not only for operating their service, but also for promoting their services, and for developing new ones:

When you upload, submit, store, send or receive content to or through our Services, you give Google (and those we work with) a worldwide license to use, host, store, reproduce, modify, create derivative works [...], communicate, publish, publicly perform, publicly display and distribute such content. The rights you grant in this license are for the limited purpose of operating, promoting, and improving our Services, and to develop new ones. This license continues even if you stop using our Services [...].

It is unclear what use of the content would fall *outside* of the scope of this license.

There is no way that you can count on the service doing today what it did yesterday, because Google reserves the right to change the service whenever they want, and then to force this change upon you:

When a Service requires or includes downloadable software, this software may update automatically on your device once a new version or feature is available.

You can't even count on the service to be there tomorrow:

We may add or remove functionalities or features, and we may suspend or stop a Service altogether.

Basically Google doesn't want to take responsibility for their service doing anything. So it won't make any promises that any of their services will do anything useful.

Other than as expressly set out in these terms or additional terms, neither Google nor its suppliers or distributors make any specific promises about the

²¹¹For example through hijacking the two-factor SMS code. See: Gibbs, "SS7 Hack Explained." or Franceschi-Bicchierai, "The SIM Hijackers."

Services. For example, we don't make any commitments about the content within the Services, the specific functions of the Services, or their reliability, availability, or ability to meet your needs.

And, it wants to make sure that the user understands that there isn't any warranties and that Google won't take the responsibility for any losses:

To the extent permitted by law, we exclude all warranties. [...] When permitted by law, Google, and Google's suppliers and distributors, will not be responsible for lost profits, revenues, or data, financial losses or indirect, special, consequential, exemplary, or punitive damages.

If, for some reason, they are still forced to pay damages, then they limit their own liability to whatever the user has paid for the services. In the case of Gmail and Google Photos for example, this payment amounts to zero (in monetary terms that is):

To the extent permitted by law, the total liability of Google, and its suppliers and distributors, for any claims under these terms, including for any implied warranties, is limited to the amount you paid us to use the Services [...].

Finally, Google reserves the right to change these terms of service at any point in time and expects the user to look at them regularly to make sure they've noticed the change. If they don't like the change, then the only option left for the user is to stop using the service.

We may modify these terms or any additional terms that apply to a Service to, for example, reflect changes to the law or changes to our Services. You should look at the terms regularly. [...] If you do not agree to the modified terms for a Service, you should discontinue your use of that Service.

My much shorter version of these terms would be: "We at Google take no responsibility for anything, and you the user have no rights. And even though we can do what we want and you can expect nothing from us, we still want to be able to change this agreement whenever we feel like it."

I am aware that much of this is standard legalese, and that some of these terms are limited by what the law allows (hence the few occasions of "to the extent permitted by law"), but I also find the way that these terms are formulated idiosyncratic for the particular relationship that we have with companies like Google. Imagine if these were the terms that you had to sign before filling up at a gas station, or when buying a laptop.

In a sense, Google can be compared to ransomware. Ransomware encrypts your digital life and gives you the decryption key as soon as you paid the required ransom in cryptocurrency, whereas Google will only allow you to continue to have access to your digital life as long as

you comply with their loaded terms.

The fact that Google can make arbitrary decisions and subject their users to their will breaks Pettit's third clause for making a choice free. Pettit uses the eyeball test (you should be able "to look one another in the eye without reason for fear or deference"²¹²) as a way to know when a free person has enough protections against arbitrary control. He has a version of the test, that he uses for our international relations, that I think is more fitting to administer to our relationship with Google and the other technology giants:

Each people in the world ought to be able to address other peoples [...] as an equal among equals. It ought not to be required to resort to the tones of a subservient subject and it ought not to be entitled to adopt the arrogant tones of a master. It ought to enjoy the capacity to frame its expectations and proposals on the assumption of having a status no lower and no higher than others and so to negotiate in a straight-talking, open manner. Each people ought to be able to pass what we might call the straight talk test.²¹³

We can't pass this straight talk test in our technological predicament. We are living an increasing part of our lives inside corporate terms of service. To the extent that we live under their arbitrary governance, we can't consider ourselves to have civic freedom.²¹⁴

²¹²Pettit, *Just Freedom*, 90.

²¹³Ibid., 181–82.

²¹⁴Rawls also understands the freedom limiting aspects of arbitrary decisions: "But if the precept of no crime without a law is violated, say by statutes, being vague and imprecise, what we are at liberty to do is likewise vague and imprecise. The boundaries of our liberty are uncertain. And to the extent that this is so, liberty is restricted by a reasonable fear of its exercise." See: Rawls, *A Theory of Justice* 210.

Part 3: What should we do about it?

In a situation where one party is more powerful than another party, there are basically three things you can do to create antipower. You can diminish the power of the first party, you can regulate the first party in such a way that there is no way for them to exercise their power, or you can empower the second party. As Pettit writes:

We may compensate for imbalances by giving the powerless protection against the resources of the powerful, by regulating the use that the powerful make of their resources, and by giving the powerless new, empowering resources of their own. We may consider the introduction of protective, regulatory, and empowering institutions.²¹⁵

For Pettit it is clear that antipower can't just come from the legal instruments with which the state operates, there is also a clear role for the various institutions inside civil society.²¹⁶ In the final part of this thesis, I will do some short speculative explorations of potential directions towards bettering our technological predicament. Consider these “plays” in the antipower playbook.

These explorations stay very close to the three core characteristics of our technological predicament. To counter the domineering scale, we need to look at ways of reducing the scale; to address data-driven appropriation, we need to reinvigorate our commons; and to deal with the problem of asymmetrical relationships with arbitrary control, we need to see how we can use technology to design equality in our relationships.²¹⁷

Reducing the scale

There are two obvious ways to reduce the scale at which our communications infrastructure operates. We can try to make the technology giants smaller (or at least stop them from getting any bigger and more dominant), or we can try and switch to a technological infrastructure that still allows us to connect at a world scale, without creating similar dependencies as in our current technological predicament.

²¹⁵Pettit, “Freedom as Antipower,” 589–90.

²¹⁶Ibid., 593.

²¹⁷It would probably be feasible to argue that many of the problems in our technological predicament have the prevailing neo-liberal form of capitalism as their root cause. This means that solutions to the problem would need to consist of finding pathways to new economic arrangements. As this thesis does not contain a capitalist critique, these explorations don't explicitly address economic systems either. However, it is glaringly obvious that it will require appropriate adjustments to our accumulation mindset for many of these ideas to be successful.

Traditional antitrust legislation tries to battle the negative effects of monopolies through looking at how market domination affects the price for the consumer. A classic antitrust measure is to bust a cartel that has artificially fixed the prices. Looking at prices becomes close to meaningless in a situation where the consumer doesn't appear to pay anything, and is—on the surface—better off using the product rather than not using the product.

The different data flows and market dominance from a user perspective don't get enough focus in decisions about antitrust. This is why the European Commission made the mistake of allowing Facebook to buy its competitor WhatsApp for 19 billion U.S. dollars.²¹⁸ Facebook told the Commission in 2014 that it would not be technically feasible to reliably automate the matching between Facebook user accounts and the accounts of WhatsApp. In August 2016, Facebook did exactly that, and eventually was fined 110 million euro for this behavior.²¹⁹

Facebook's acquisition of WhatsApp is part of a larger pattern. Big giants like Google, Microsoft, and Facebook prefer to buy up smaller competing companies who are delivering an innovative product.²²⁰ But if these companies refuse to be bought, they will just make a blatant copy of their functionality. Some people call it a "kill-zone" around the internet giants.²²¹ For example, Facebook was able to buy Instagram, but couldn't get its hands on Snapchat. So it copied most of Snapchat's features into Instagram.²²² Facebook has even bought Onavo, an app that monitors what people are doing on their phone. It uses the aggregated data of the millions of users of the app, to see what services are popular, in order to snap them up before they get too big and endanger the size of Facebook's user base.²²³²²⁴

The Economist therefore recommends that antitrust authorities start taking a different approach. Instead of using just size to determine whether to intervene, "they now need to take into account the extent of firms' data assets when assessing the impact of deals. The purchase price could also be a signal that an incumbent is buying a nascent threat."²²⁵

A more technical approach than making use of antitrust law, is to work on alternatives to the big companies. One of the incredible things about the internet is that the network facilitates peer-to-peer interactions. It is possible to have a direct connection between two internet enabled devices (for example two smartphones) and have an encrypted set of communications

²¹⁸Rushe, "WhatsApp."

²¹⁹"Commission Fines Facebook €110 Million for Providing Misleading Information About WhatsApp Takeover."

²²⁰Alphabet, Amazon, Apple, Facebook, and Microsoft together spent \$31.6bn on acquisitions in 2017. See: Economist, "American Tech Giants Are Making Life Tough for Startups."

²²¹Ibid.

²²²Pierce, "Facebook Has All of Snapchat's Best Features Now."

²²³Griffith, "Will Facebook Kill All Future Facebooks?"

²²⁴Google is in a similar information position, through owning the Google app store, and through its Chrome browser.

²²⁵Economist, "The World's Most Valuable Resource Is No Longer Oil, but Data."

data flow between them. This allows for a typology of different ways to federate or decentralize technological infrastructure. The following are just three examples of technology projects that reduce scale, and therefore reduce domination:

- Mastodon²²⁶ is an open source social network allowing users to post short messages, pictures, and videos in a similar way to Facebook and Twitter. Unlike other social networks it is fully decentralized. There is no one single company or server that contains all the messages. Instead, different “instances” of Mastodon have a way of talking to one another. Each instance can have their own rules about what type of content it allows and which people they will give accounts on their system. Users within an instance can follow each other, but it is also possible to follow people who have their home base at another instance.
- Briar²²⁷ is a secure messaging app that allows peer-to-peer encrypted messaging and forums. It breaks with the normal messaging paradigm which relies on a central server to receive and deliver messages. Briar only delivers messages when both parties have an internet connection at the same time. It can do this locally using a Bluetooth connection or a Wi-Fi network, but it can also use an internet connection. In the latter case, it will route the traffic over Tor in order to ensure its anonymity and to hide the user’s location. Connections on Briar are made by being together physically and exchanging keys. All of these design choices make Briar very resistant to both surveillance and censorship. The app can even keep local communication flowing during internet blackouts.
- The Dat Project²²⁸ is host of the Dat Protocol, a peer-to-peer data sharing protocol that allows for distributive syncing. With Dat’s network users can store data wherever they want (with most data being stored at multiple locations). Dat keeps a history of how a file has changed, facilitating collaboration and easy reproducibility. Users can easily replicate a remote Dat repository and subscribe to live changes. Network traffic is encrypted, and it is possible to create your own private data sharing networks.²²⁹

It is also possible to resist scale by making your own websites and tools. I have built a few sites myself which I call “hyperpersonal microsites”, because they mainly have an audience of one (even though they are public) and serve a single purpose. In this way, I have replaced my use of Amazon’s Goodreads, a social network for readers, with my own website for storing what books I have read, which ones I still want to read, and my book reviews.²³⁰ Rather than feeding large corporations with data about my reading habits, I now make use of their

²²⁶“Mastodon.”

²²⁷“How It Works.”

²²⁸“Dat Project - A Distributed Data Community.”

²²⁹Ogden et al., “Dat - Distributed Dataset Synchronization And Versioning.”

²³⁰De Zwart, “Hans de Zwart’s Books.”

application programming interfaces (APIs) for my own purposes. A similar project, is a small website that allows me to answer my main mapping need (how to get to somewhere in Amsterdam on my bike, from my home or from my work) in a quicker way than with Google Maps, while relying on the communal data of OpenStreetMap for the routing.²³¹

Within all these projects lies the danger of technological elitism: they require a lot of knowledge to get going and to operate. But they do make it clear that escaping from the dominating scale of big technology companies is only possible by making very conscious long term technology choices. Democratizing access to the internet, to coding skills, and to the hardware that is necessary to make things for yourself, should therefore be paramount.

Reinvigorating the commons

The P2P Foundation has put a lot of effort into conceptualizing the commons which, according to them, can be understood from at least four different perspectives:

1. Collectively managed resources, both material and immaterial, which need protection and require a lot of knowledge and know-how.
2. Social processes that foster and deepen thriving relationships. These form part of complex socio-ecological systems which must be consistently stewarded, reproduced, protected and expanded through commoning.
3. A new mode of production focused on new productive logics and processes.
4. A paradigm shift, that sees commons and the act of commoning as a worldview.²³²

It is important to realize that “the Commons is neither the resource, the community that gathers around it, nor the protocols for its stewardship, but the dynamic interaction between all these elements.”²³³ The P2P Foundation sees peer-to-peer relations in their non-hierarchical and non-coercive form as one of the “enabling capacities for actions. [Peer-to-peer] facilitates the act of ‘commoning,’ as it builds capacities to contribute to the creation of maintenance of any shared and co-managed resource (a commons).”²³⁴ An important example of a commons in the context of this thesis, is Wikipedia.

One obvious way to reinvigorate the commons, is to explicitly invest into commoning projects like Wikipedia and OpenStreetMap, and also to start seeing them as commons,

²³¹De Zwart, “Hans Fietst.”

²³²“Commons Transition and P2P,” 5, emphasis removed.

²³³Ibid., 5.

²³⁴Ibid., 10.

rather than as a simple free resource. But doing this wouldn't necessarily intervene directly into the data-driven appropriation of the accumulators and their abuse of the informational and attentional commons. There is a dearth of academic work in this space,²³⁵ so the following couple of ideas are necessarily very rough and underdeveloped.

A first change would be to start thinking about ecosystem (or collective) rights in addition to individual rights. Currently, most data protection law tries to intervene at the individual level, as it describes individual rights. This means that it can't address collective problems from processes that don't deal with personal data (for example what Vodafone does with mobility data, as mentioned in the introduction). We need to start thinking about what it means for society if we allow private companies to capture all of the externalities of the use of their services, and then sell that information about the world back to the public.

We could also consider a flat out ban on the appropriation of data (as earlier defined in this thesis) for private purposes. This would forbid private companies from collecting and using data against people's will or without their knowledge. It would be important to combine these rules with very strict purpose limitations: data that is collected (with knowledge and free permission) for one purpose cannot be used for another purpose. If this seems too radical, then an alternative would be to require private companies to open up the non-personal (or fully anonymized) data they have gathered, and make it available inside a data commons with open licenses.

It is interesting to think about what would happen if we were to take a "right to be left alone" seriously, and work towards an attentional commons. Someone who saw the importance of this was Gilberto Kassab, the mayor of São Paulo. In 2007, as part of his Clean City Law, he put into effect a near complete ban of outdoor advertising in his city: "The Clean City Law came from a necessity to combat pollution [...] pollution of water, sound, air, and the visual. We decided that we should start combating pollution with the most conspicuous sector – visual pollution."²³⁶ The results were interesting: it encouraged companies to reassess their advertising campaigns and find new and creative ways to engage with their customers. All without covering up the architecture of the city.²³⁷ It is hard to imagine the virtual analogy to the Clean City Law, but approaching our virtual spaces from the perspective of abating cognitive pollution certainly could help.

A final idea to stop data-driven appropriation, is to require of the technology giants that they provide access to their data through open standards and through open application pro-

²³⁵The most interesting work is probably done at the Data & Society research institute in New York, which focuses on the social and cultural issues arising from data-centric and automated technologies. See: "What We Do."

²³⁶"São Paulo."

²³⁷Goodson, "No Billboards, No Outdoor Advertising?"

gramming interfaces (APIs). Privacy technology specialist Jaap-Henk Hoepman has written about this idea on his blog.²³⁸ Hoepman starts by explaining how email is an open standard, which means that you can exchange emails with other people, regardless of what program they use to access their email. This compatibility isn't the case with messages between Apple's iMessage, Instagram, Skype, WhatsApp and other messaging clients. According to Hoepman, this is as if Outlook users would only be able to email with other Outlook users, or if you could only text from your Nokia phone to other people with a Nokia phone, or only to people with the same mobile service provider. Forcing the use of open standards and open APIs should allow Apple to find a way to let iMessage talk to WhatsApp and might even allow truly open alternatives to ride the coat-tails of the network effects that are enabling the technology giants.

Equality in relationships

When trying to battle the asymmetry in relationships with the technology giants, it is important to find a way to break up the user lock-in that affects our relationship with companies like Google and Facebook. Two workable ways that this can be done are through breaking up the lock-in with a requirement for data portability, and through never stepping into the lock-in by using free, instead of proprietary, software.

Data portability is the idea that it should be possible to transfer your data from one service to another, preferably in an automated fashion. Europe's General Data Protection Regulation (GDPR) defines data portability as an explicit right for all the people who are residing in the European Union. The regulation defines the right to data portability as follows:

The data subject shall have the right to receive the personal data concerning him or her, which he or she has provided to a controller, in a structured, commonly used and machine-readable format and have the right to transmit those data to another controller without hindrance from the controller to which the personal data have been provided [...]. In exercising his or her right to data portability [...] the data subject shall have the right to have the personal data transmitted directly from one controller to another, where technically feasible.²³⁹

Data portability has some challenges when it comes to privacy (if I want to move all my social networking contacts from service A to service B, do all my contacts want service B to now know about their existence?), and this is why the EU has added to the right to data

²³⁸Hoepman, "Doorbreek Monopolies Met Open Standaarden En API's."

²³⁹"General Data Protection Regulation" Article 20.

portability that it “shall not adversely affect the rights and freedoms of others.”²⁴⁰ But these challenges are in no way insurmountable. It is important that we urgently start expecting a lot more maturity from the likes of Google and Facebook in making this right a concrete reality. The GDPR’s data regime already has had an effect:²⁴¹ Facebook, Google, Twitter and Microsoft have recently launched the Data Transfer Project, which aims to provide users with “the ability to initiate a direct transfer of their data into and out of any participating provider.”²⁴²

Rather than making use of data portability, it is also possible to never step into a locked-in situation. This is enabled through what is called “free software”. This software uses the term free as in liberty, it isn’t about price. It guarantees freedom through a legal license which was initially developed by Richard Stallman.²⁴³ Stallman wants anybody who uses software to have what he calls “the four freedoms” (which he purposefully starts counting at zero, like any computer engineer would do):

0. The freedom to run the program as you wish, for any purpose.
 1. The freedom to study how the program works, and change it so it does your computing as you wish.
 2. The freedom to redistribute copies so you can help others.
 3. The freedom to distribute copies of your modified versions to others. By doing this you can give the whole community a chance to benefit from your changes.²⁴⁴

Stallman sees guaranteeing these freedoms as a moral imperative for developers of software. Any free software licence²⁴⁵ protects the user from an asymmetrical relationship with the creators of the software. There can’t be a vendor lock-in because there are no barriers to entry for providing services around a free software product. Through using free software you can inoculate yourself against domineering technology companies.

It behooves the state to promote the use of free software. According to Stallman the state should only use free software for their own computing, should only teach the use of free software in schools, should never require its citizens to use non-free programs to access state services, and should incentivize and patron the development of free software. With these measures the state can recover control over its computing, and help citizens, businesses and

²⁴⁰Ibid. Article 20.

²⁴¹Claburn, “Facebook, Google, Microsoft, Twitter Make It Easier to Download Your Info and Upload to, Er, Facebook, Google, Microsoft, Twitter Etc.”

²⁴²“Data Transfer Project Overview and Fundamentals.”

²⁴³Stallman also has a radical proposal to keep our personal data safe: create laws that stop data appropriation. See: Stallman, “A Radical Proposal to Keep Your Personal Data Safe.”

²⁴⁴“What Is Free Software?”

²⁴⁵See “The Open Source Definition.” for a definition of what make a license free.

organizations to do the same.²⁴⁶

²⁴⁶Stallman, “Measures Governments Can Use to Promote Free Software, And Why It Is Their Duty to Do so.”

Acknowledgements

Writing this thesis would not have been possible without the two years of support and boundless patience and flexibility from my partner. It is an incredible privilege to share a life with someone who emanates that much love and joy.

I am also very thankful to my mother and her life partner for offering me a place to write. It allowed for the necessary solitude, while also giving me the opportunity to discuss my progress during the three—fully catered—meals a day. I am grateful for the pleasure of working with my wonderful colleagues at Bits of Freedom every single working day. The countless discussions with them, and with people from the broader digital rights movement, have certainly sharpened my thinking.

In academia, I would like to explicitly thank Beate Roessler for her no-nonsense approach to supervision. She managed to always increase clarity, both for the process and for the content. Thomas Nys for his willingness to be the second reader, and Gijs van Donselaar for introducing me to Pettit’s thinking, and for sharpening my Bachelor’s thesis. Special thanks go out to Philip Pettit, for making the time to have a conversation with me in Prague to share his thoughts on republicanism and our technological predicament.

The software stack that I used for the writing of this thesis was completely free from domination. I want to thank the creators of all the free software that enabled me to write with a conscience. I can’t do justice to the layers of work upon other people’s work that allow for my computer to function as it does, but I do want to at least acknowledge the creators and maintainers of GNU/Linux and Ubuntu (for providing the core of my operating system), i3 (a tiling window manager), Firefox (my browser of choice), Vim (the text editor allowing me to “edit text at the speed of thought”²⁴⁷), Zotero (for storing my references), and Markdown, LaTeX and Pandoc (enabling the workflow from a text file to a beautifully typeset PDF).

Now that this is done, I look forward to putting more energy in getting us out of our technological predicament, and in helping to build a just and free alternative technological infrastructure.

Yours in struggle,

Hans de Zwart

Amsterdam, August 2018

²⁴⁷Neil, *Practical Vim*.

Bibliography

- “About Seth.” *Seth Stephens-Davidowitz*. Accessed July 8, 2018. <http://sethsd.com/>.
- Alli, Kabir. “YO OOOO LOOK AT THIS.” Tweet. @IBeKabir, June 2016. <https://twitter.com/iBeKabir/status/740005897930452992>.
- Anderson, Chris. “The End of Theory: The Data Deluge Makes the Scientific Method Obsolete.” *Wired*, June 2008. <https://www.wired.com/2008/06/pb-theory/>.
- Angwin, Julia, and Hannes Grassegger. “Facebook’s Secret Censorship Rules Protect White Men From Hate Speech But Not Black Children.” Text/html. *ProPublica*, June 2017. <https://www.propublica.org/article/facebook-hate-speech-censorship-internal-documents-algorithms>.
- “Annual Report - Google Diversity.” Accessed July 21, 2018. <https://diversity.google/annual-report/>.
- Austin, Evelyn. “Women on Waves’ Three YouTube Suspensions This Year Show yet Again That We Can’t Let Internet Companies Police Our Speech.” *Bits of Freedom*, June 2018. <https://www.bof.nl/2018/06/28/women-on-waves-three-youtube-suspensions-this-year-show-yet-again-that-we-cant-let-internet-companies-police-our-speech/>.
- Ayers, Phoebe. “YouTube Should Probably Run Some A/B Tests with the Crew at @WikiResearch First.” Tweet. @Phoebe_ayers, March 2018. https://twitter.com/phoebe_ayers/status/973742197857284096.
- Balkan, Aral. “Encouraging Individual Sovereignty and a Healthy Commons,” February 2017. <https://ar.al/notes/encouraging-individual-sovereignty-and-a-healthy-commons/>.
- . “The Nature of the Self in the Digital Age,” March 2016. <https://2018.ar.al/notes/the-nature-of-the-self-in-the-digital-age/>.
- Barocas, Solon, and Helen Nissenbaum. “Big Data’s End Run Around Procedural Privacy Protections.” *Communications of the ACM* 57, no. 11 (October 2014): 31–33. doi:10.1145/2668897.
- Berlin, Isaiah. “Two Concepts of Liberty.” In *Liberty: Incorporating Four Essays on Liberty*, edited by Henry Hardy, 166–217. Oxford: Oxford University Press, 2002.
- Bickert, Monika. “Publishing Our Internal Enforcement Guidelines and Expanding Our Appeals Process.” *Facebook Newsroom*, April 2018. <https://newsroom.fb.com/news/2018/04/publishing-our-internal-enforcement-guidelines-and-expanding-our-appeals-process/>.

[8/04/comprehensive-community-standards/](#).

Bliss, Laura. "The Real Problem With 'Areas of Interest' on Google Maps." *CityLab*, August 2016. <http://www.citylab.com/design/2016/08/google-maps-areas-of-interest/493670/>.

Borgers, Eddie. "Marktaandeelen Zoekmachines Q1 2018." *Pure*, April 2018. <https://www.pure-im.nl/blog/marktaandeelen-zoekmachines-q1-2018/>.

Boyle, James. "The Second Enclosure Movement." *Renewal* 15, no. 4 (2007): 17–24.

Bridle, James. "Something Is Wrong on the Internet." *James Bridle*, November 2017. <https://medium.com/@jamesbridle/something-is-wrong-on-the-internet-c39c471271d2>.

Brouwer, Bree. "YouTube Now Gets Over 400 Hours Of Content Uploaded Every Minute." *Tubefilter*, July 2015. <https://www.tubefilter.com/2015/07/26/youtube-400-hours-content-every-minute/>.

Cadwalladr, Carole, and Emma Graham-Harrison. "Revealed: 50 Million Facebook Profiles Harvested for Cambridge Analytica in Major Data Breach." *The Guardian*, March 2018. <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>.

Calo, Ryan, and Alex Rosenblat. "The Taking Economy: Uber, Information, and Power." *Columbia Law Review* 117 (March 2017). doi:10.2139/ssrn.2929643.

Chappell, Bill. "Google Maps Displays Crimean Border Differently In Russia, U.S." *NPR*, April 2014. <https://www.npr.org/sections/thetwo-way/2014/04/12/302337754/google-maps-displays-crimean-border-differently-in-russia-u-s>.

"Choose Your Audience." *Facebook Business*. Accessed June 23, 2018. <https://www.facebook.com/business/products/ads/ad-targeting>.

Claburn, Thomas. "Facebook, Google, Microsoft, Twitter Make It Easier to Download Your Info and Upload to, Er, Facebook, Google, Microsoft, Twitter Etc." *The Register*, July 2018. https://www.theregister.co.uk/2018/07/20/facebook_google_microsoft_twitter_data/.

"Commission Fines Facebook €110 Million for Providing Misleading Information About WhatsApp Takeover." *European Commission Press Releases*, May 2017. http://europa.eu/rapid/press-release_IP-17-1369_en.htm.

"Commons Transition and P2P: A Primer." Transnational Institute, March 2017. https://www.tni.org/files/publication-downloads/commons_transition_and_p2p_primer_v9.pdf.

Conger, Kate, and Dell Cameron. "Google Is Helping the Pentagon Build AI for Drones."

Gizmodo, June 2018. <https://gizmodo.com/google-is-helping-the-pentagon-build-ai-for-drones-1823464533>.

Court of Justice. “Google Spain SL and Google Inc. V Agencia Española de Protección de Datos (AEPD) and Mario Costeja González,” May 2014. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:62012CJ0131>.

Crawford, Matthew B. *The World Beyond Your Head: On Becoming an Individual in an Age of Distraction*. New York: Farrar, Straus; Giroux, 2015.

Curran, Dylan. “Are You Ready? This Is All the Data Facebook and Google Have on You.” *The Guardian*, March 2018. <http://www.theguardian.com/commentisfree/2018/mar/28/all-the-data-facebook-google-has-on-you-privacy>.

“Dat Project - A Distributed Data Community.” *Dat Project*. Accessed August 13, 2018. datproject.org/.

“Data Drives All That We Do.” *Cambridge Analytica*. Accessed June 24, 2018. <https://cambridgeanalytica.org/>.

“Data Transfer Project Overview and Fundamentals,” July 2018. <https://datatransferproject.dev/dtp-overview.pdf>.

“Data-Driven Campaigns.” *CA Political*. Accessed June 24, 2018. <https://ca-political.com/>.

De Zwart, Hans. “Demystifying the Algorithm.” *Hans de Zwart*, June 2015. <https://medium.com/@hansdezward/demystifying-the-algorithm-6e52729e3576>.

———. “Facebook Is Gemaakt Voor Etnisch Profileren.” *De Volkskrant*, June 2016. <https://www.volkskrant.nl/g-be1071d6>.

———. “Google Wijst Me de Weg, Maar Niet Altijd de Kortste.” *NRC*, August 2015. <https://www.nrc.nl/nieuws/2015/08/15/google-wijst-me-de-weg-maar-niet-altijd-de-kortst-1523170-a861546>.

———. “Hans de Zwart’s Books.” Accessed August 13, 2018. <https://books.hansdezward.nl/>.

———. “Hans Fietst.” Accessed August 13, 2018. <https://hans.fietst.org/>.

———. “Liberty, Technology and Democracy.” Amsterdam: University of Amsterdam, August 2017. <https://blog.hansdezward.nl/2017/09/03/liberty-technology-and-democracy/>.

———. “Medium Massage – Writings by Hans de Zwart.” Accessed July 21, 2018. <https://blog.hansdezward.nl/>.

———. “Miljardenbedrijf Google Geeft Geen Cent Om de Waarheid.” *NRC*, August 2018.

<https://www.nrc.nl/nieuws/2018/08/08/miljardenbedrijf-google-geeft-geen-cent-om-de-waarheid-a1612498>.

“DeepMind Health.” *DeepMind*. Accessed June 24, 2018. <https://deepmind.com/applied/deepmind-health/>.

“Definition of Appropriate in the Merriam Webster Dictionary.” *Merriam Webster*. Accessed June 24, 2018. <https://www.merriam-webster.com/dictionary/appropriate>.

“Definition of Appropriate in the Oxford Dictionary.” *Oxford Dictionaries*. Accessed June 24, 2018. <https://en.oxforddictionaries.com/definition/appropriate>.

Dinzeo, Maria. “Google Ducks Gmail Captcha Class Action.” *Courthouse News Service*, February 2016. <https://www.courthousenews.com/google-ducks-gmail-captcha-class-action/>.

“Donald J. Trump for President.” *CA Political*. Accessed June 24, 2018. <https://ca-political.com/casestudies/casestudydonaldjtrumpforpresident2016>.

Economist, The. “American Tech Giants Are Making Life Tough for Startups.” *The Economist*, June 2018. <https://www.economist.com/business/2018/06/02/american-tech-giants-are-making-life-tough-for-startups>.

———. “The World’s Most Valuable Resource Is No Longer Oil, but Data.” *The Economist*, May 2017. <http://www.economist.com/node/21721656>.

Ehrlich, Jamie. “GOP Senator Says He Is Alive Amid Google Searches Suggesting He Is Dead.” *CNN*, July 2018. <https://www.cnn.com/2018/07/24/politics/orrin-hatch-alive-google-search-dead/index.html>.

Epstein, Robert. “How Google Could Rig the 2016 Election.” *POLITICO Magazine*, August 2015. <https://www.politico.com/magazine/story/2015/08/how-google-could-rig-the-2016-election-121548.html>.

Epstein, Robert, and Ronald E. Robertson. “The Search Engine Manipulation Effect (SEME) and Its Possible Impact on the Outcomes of Elections.” *Proceedings of the National Academy of Sciences* 112, no. 33 (August 2015): E4512–E4521. doi:10.1073/pnas.1419828112.

Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin’s Press, 2018.

“Facebook Ads Targeting Comprehensive List.” *Two Wheels Marketing*, January 2018. <https://www.twowheelsmarketing.com/facebook-ads-targeting-comprehensive-list/>

[tps://twowheelsmarketing.com/blog/facebook-ads-targeting-options-list/](https://twowheelsmarketing.com/blog/facebook-ads-targeting-options-list/).

Falcone, John. "Amazon Backtracks, Will Offer \$15 Opt-Out for Ads on Kindle Fire Tablets." *CNET*, September 2012. <https://www.cnet.com/news/amazon-backtracks-will-offer-15-opt-out-for-ads-on-kindle-fire-tablets/>.

Farokhmanesh, Megan. "YouTube Didn't Tell Wikipedia About Its Plans for Wikipedia." *The Verge*, March 2018. <https://www.theverge.com/2018/3/14/17120918/youtube-wikipedia-conspiracy-theory-partnerships-sxsw>.

Franceschi-Bicchierai, Lorenzo. "The SIM Hijackers." *Motherboard*, July 2018. https://motherboard.vice.com/en_us/article/vbqax3/hackers-sim-swapping-steal-phone-numbers-instagram-bitcoin.

Gellman, Robert. "Disintermediation and the Internet." *Government Information Quarterly* 13, no. 1 (January 1996): 1–8. doi:10.1016/S0740-624X(96)90002-7.

"General Data Protection Regulation," May 2018. <https://eur-lex.europa.eu/eli/reg/2016/679/oj>.

Gibbs, Samuel. "SS7 Hack Explained: What Can You Do About It?" *The Guardian*, April 2016. <http://www.theguardian.com/technology/2016/apr/19/ss7-hack-explained-mobile-phone-vulnerability-snooping-texts-calls>.

Goodrow, Cristos. "You Know What's Cool? A Billion Hours." *Official YouTube Blog*, February 2017. <https://youtube.googleblog.com/2017/02/you-know-whats-cool-billion-hours.html>.

Goodson, Scott. "No Billboards, No Outdoor Advertising? What Next?" *Forbes*, January 2012. <https://www.forbes.com/sites/marketshare/2012/01/04/no-billboards-no-outdoor-advertising-what-next/>.

"Google Search Statistics." *Internet Live Stats*. Accessed July 8, 2018. <http://www.internetlivestats.com/google-search-statistics/>.

"Google Terms of Service," October 2017. <https://policies.google.com/terms>.

"Google Trends." *Google Trends*. Accessed July 8, 2018. <https://trends.google.com>.

Gregory, Karen. "Big Data, Like Soylent Green, Is Made of People." *Digital Labor Working Group*, November 2014. <https://digitallabor.commons.gc.cuny.edu/2014/11/05/big-data-like-soylent-green-is-made-of-people/>.

Griffith, Erin. "Will Facebook Kill All Future Facebooks?" *Wired*, October 2017. <https://www.wired.com/2017/10/facebook-future/>.

[//www.wired.com/story/facebooks-aggressive-moves-on-startups-threaten-innovation/](http://www.wired.com/story/facebooks-aggressive-moves-on-startups-threaten-innovation/).

Harris, David L. “Massachusetts Woman’s Lawsuit Accuses Google of Using Free Labor to Transcribe Books, Newspapers.” *Boston Business Journal*, January 2015. <https://www.bizjournals.com/boston/blog/techflash/2015/01/massachusetts-womans-lawsuit-accuses-google-of.html>.

Hern, Alex. “Facebook Protects Far-Right Activists Even After Rule Breaches.” *The Guardian*, July 2018. <https://www.theguardian.com/technology/2018/jul/17/facebook-protects-far-right-activists-even-after-rule-breaches-dispatches-c4>.

Hess, Charlotte, and Elinor Ostrom. “Introduction: An Overview of the Knowledge Commons.” In *Understanding Knowledge as a Commons*, 3–26. Cambridge, Massachusetts: The MIT Press, 2007.

Hirsch Ballin, Ernst, Dennis Broeders, Erik Schrijvers, Bart van der Sloot, Rosamunde van Brakel, and Josta de Hoog. “Big Data in Een Vrije En Veilige Samenleving.” Den Haag: Wetenschappelijke Raad voor het Regeringsbeleid/Amsterdam University Press, April 2016.

Hobbes, Thomas. *Leviathan*. Edited by Richard Tuck. Cambridge: Cambridge University Press, 1996.

Hoepman, Jaap-Henk. “Doorbreek Monopolies Met Open Standaarden En API’s,” February 2018. <https://blog.xot.nl/2018/02/15/doorbreek-monopolies-met-open-standaarden-en-apis/>.

“How Google Retains Data We Collect.” *Google Privacy & Terms*. Accessed June 23, 2018. <https://policies.google.com/technologies/retention>.

“How It Works.” *Briar*. Accessed August 13, 2018. <https://briarproject.org/how-it-works.html>.

Ibsen, Henrik. *A Doll’s House*. Gloucester: Dodo Press, 2015.

“Internet: Toegang, Gebruik En Faciliteiten.” *Centraal Bureau Voor de Statistiek - StatLine*. Accessed June 23, 2018. <https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83429NED/table?ts=1529749546372>.

Jagadish, H. V., Johannes Gehrke, Alexandros Labrinidis, Yannis Papakonstantinou, Jignesh M. Patel, Raghu Ramakrishnan, and Cyrus Shahabi. “Big Data and Its Technical Challenges.” *Communications of the ACM* 57, no. 7 (July 2014): 86–94. doi:10.1145/2611567.

Kamona, Bonnie. “I Saw a Tweet Saying ‘Google Unprofessional Hairstyles for Work.’”

Tweet. @HereroRocher, April 2016. <https://twitter.com/HereroRocher/status/717457819864272896>.

Kant, Immanuel. *Notes and Fragments*. Cambridge: Cambridge University Press, 2005.

Kreiken, Floris. “Humanitair-Vrijheids-Vrede-Mensenrechten-Project-Facebook.” *Bits of Freedom*, August 2014. <https://www.bof.nl/2014/08/05/humanitair-vrijheids-vrede-mensenrechten-project-facebook/>.

Kreling, Tom, Huib Modderkolk, and Maartje Duin. “De Hel Achter de Façade van Facebook.” *Volkskrant*, April 2018. <https://www.volkskrant.nl/gs-b884f37c5>.

Levy, Steven. “How Google’s Algorithm Rules the Web.” *Wired*, February 2010. https://www.wired.com/2010/02/ff_google_algorithm/.

Li, Mark, and Zhou Bailang. “Discover the Action Around You with the Updated Google Maps.” *The Keyword*, July 2016. <https://www.blog.google/products/maps/discover-action-around-you-with-updated/>.

“List of Public Corporations by Market Capitalization.” *Wikipedia*, June 2018. https://en.wikipedia.org/w/index.php?title=List_of_public_corporations_by_market_capitalization&oldid=846735785.

Madrigal, Alexis C. “How Google Builds Its Maps—and What It Means for the Future of Everything.” *The Atlantic*, September 2012. <https://www.theatlantic.com/technology/archive/2012/09/how-google-builds-its-maps-and-what-it-means-for-the-future-of-everything/261913/>.

“Mastodon.” Accessed August 13, 2018. <https://joinmastodon.org/>.

Matsakis, Louise. “Don’t Ask Wikipedia to Cure the Internet.” *Wired*, March 2018. <https://www.wired.com/story/youtube-wikipedia-content-moderation-internet/>.

———. “YouTube Will Link Directly to Wikipedia to Fight Conspiracy Theories.” *Wired*, March 2018. <https://www.wired.com/story/youtube-will-link-directly-to-wikipedia-to-fight-conspiracies/>.

Miller, Ron. “Cheaper Sensors Will Fuel The Age Of Smart Everything.” *TechCrunch*, March 2015. <https://techcrunch.com/2015/03/10/cheaper-sensors-will-fuel-the-age-of-smart-everything/>.

Murphy, Mike, and Akshat Rathi. “All of Google’s—er, Alphabet’s—companies and Products from A to Z.” *Quartz*, August 2015. <https://qz.com/476460/here-are-all-the-alphabet-fo>

[rmerly-google-companies-and-products-from-a-to-z/](#).

Neil, Drew. *Practical Vim: Edit Text at the Speed of Thought*. Pragmatic Bookshelf, 2015.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: NYU Press, 2018.

O’Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York: Broadway Books, 2016.

Ogden, Maxwell, Karissa McKelvey, Matthias Buus Madsen, and Code for Science. “Dat - Distributed Dataset Synchronization And Versioning,” May 2017. doi:[10.31219/osf.io/nsv2c](https://doi.org/10.31219/osf.io/nsv2c).

Ohm, Paul. “Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization.” *UCLA Law Review* 57 (2009): 1701–78.

“Our \$1 Billion Commitment to Create More Opportunity for Everyone.” *Google.org*. Accessed August 4, 2018. <https://www.google.org/billion-commitment-to-create-more-opportunity/>.

“Our Company.” *Google*. Accessed August 4, 2018. [//www.google.com/about/our-company/](https://www.google.com/about/our-company/).

“Our Mission.” *Internet.org*. Accessed August 4, 2018. <https://info.internet.org/en/mission/>.

“Our Society Is Being Hijacked by Technology.” *Center for Humane Technology*. Accessed August 7, 2018. <http://humanetech.com/problem/>.

Page, Larry. “G Is for Google.” *Alphabet*. Accessed July 8, 2018. <https://abc.xyz/>.

Perez de Acha, Gisela. “Our Naked Selves as Data – Gender and Consent in Search Engines,” April 2018. <https://www.youtube.com/watch?v=iA2wWAWJWyk>.

Pettit, Philip. “Freedom as Antipower.” *Ethics* 106, no. 3 (1996): 576–604.

———. *Just Freedom: A Moral Compass for a Complex World*. W. W. Norton & Company, 2014.

Pichai, Sundar. “AI at Google: Our Principles.” *Google*, June 2018. <https://www.blog.google/technology/ai/ai-principles/>.

Pierce, David. “Facebook Has All of Snapchat’s Best Features Now.” *Wired*, March 2017. <https://www.wired.com/2017/03/facebook-camera-means-snapping-officially-future/>.

“Predictive Policing Software.” *PredPol*. Accessed June 23, 2018. <http://www.predpol.co>

m/.

Rawls, John. *A Theory of Justice*. Revised Edition. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1999.

———. “The Priority of Right and Ideas of the Good.” *Philosophy & Public Affairs* 17, no. 4 (1988): 251–76. <http://www.jstor.org.proxy.uba.uva.nl:2048/stable/2265400>.

“reCAPTCHA.” *Wikipedia*, July 2018. <https://en.wikipedia.org/w/index.php?title=ReCAPTCHA&oldid=850679518>.

“reCAPTCHA - Creation of Value.” Accessed July 7, 2018. <https://www.google.com/recaptcha/intro/v3beta.html#creation-of-value>.

Rogers, Simon. “Data Are or Data Is? The Singular V Plural Debate.” *The Guardian*, July 2012. <http://www.theguardian.com/news/datablog/2010/jul/16/data-plural-singular>.

Ruane, Laura. “Signs of Our Times: Airport Ads Are Big Business.” *USA Today*, September 2013. <https://www.usatoday.com/story/money/business/2013/09/22/airport-ads-big-business/2842833/>.

Rushe, Dominic. “WhatsApp: Facebook Acquires Messaging Service in \$19bn Deal.” *The Guardian*, February 2014. <http://www.theguardian.com/technology/2014/feb/19/facebook-buys-whatsapp-16bn-deal>.

Safian, Robert. “Mark Zuckerberg On Fake News, Free Speech, And What Drives Facebook.” *Fast Company*, April 2017. <https://www.fastcompany.com/40397297/mark-zuckerberg-on-fake-news-free-speech-and-what-drives-facebook>.

“São Paulo: A City Without Ads.” *Adbusters*, August 2007. <https://www.adbusters.org/article/sao-paulo-a-city-without-ads/>.

Scahill, Jeremy. “The Assassination Complex.” *The Intercept*, October 2015. <https://theintercept.com/drone-papers/the-assassination-complex/>.

Schüll, Natasha Dow. *Addiction by Design: Machine Gambling in Las Vegas*. Princeton: Princeton University Press, 2012.

“Search Engine Market Share.” *NetMarketShare*. Accessed July 8, 2018. <https://hdez.nl/search-engine-market-share>.

Shane, Scott, and Daisuke Wakabayashi. “‘The Business of War’: Google Employees Protest Work for the Pentagon.” *The New York Times*, April 2018. <https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html>.

Singhal, Amit. “A Flawed Elections Conspiracy Theory.” *POLITICO Magazine*, August 2016.

<https://www.politico.com/magazine/story/2015/08/google-2016-election-121766.html>.

Smith, Kit. “39 Fascinating and Incredible YouTube Statistics.” *Brandwatch*, April 2018. <https://www.brandwatch.com/blog/39-youtube-stats/>.

“Sociale Netwerken Dagelijks Gebruik Vs. App Geïnstalleerd Nederland.” *Marketingfacts*, July 2017. <https://www.marketingfacts.nl/statistieken/detail/sociale-netwerken-dagelijks-gebruik-vs.-app-geinstalleerd-nederland>.

Stallman, Richard. “A Radical Proposal to Keep Your Personal Data Safe.” *The Guardian*, April 2018. <http://www.theguardian.com/commentisfree/2018/apr/03/facebook-abusing-data-law-privacy-big-tech-surveillance>.

———. “Measures Governments Can Use to Promote Free Software, And Why It Is Their Duty to Do so.” *GNU Project - Free Software Foundation*, January 2018. <https://www.gnu.org/philosophy/government-free-software.html>.

Stephens-Davidowitz, Seth. *Everybody Lies*. London: Bloomsbury Publishing, 2017.

———. “The Cost of Racial Animus on a Black Candidate: Evidence Using Google Search Data.” *Journal of Public Economics* 118 (October 2014): 26–40. <http://www.sciencedirect.com/science/article/pii/S0047272714000929>.

Sterling, Bruce. “Science Column #5 ‘Internet.’” *The Magazine of Fantasy and Science Fiction*, February 1993. <https://www.internetsociety.org/internet/history-internet/short-history-of-the-internet/>.

“Text of Creative Commons Attribution-ShareAlike 3.0 Unported License.” *Wikipedia*, April 2018. https://en.wikipedia.org/w/index.php?title=Wikipedia:Text_of_Creative_Commons_Attribution-ShareAlike_3.0_Unported_License&oldid=836681041.

“The Open Source Definition.” *Open Source Initiative*, March 2007. <https://opensource.org/osd>.

Thompson, Ben. “Aggregation Theory.” *Stratechery by Ben Thompson*, July 2015. <https://stratechery.com/2015/aggregation-theory/>.

———. “Antitrust and Aggregation.” *Stratechery by Ben Thompson*, April 2016. <https://stratechery.com/2016/antitrust-and-aggregation/>.

Thompson, Clive. “For Certain Tasks, the Cortex Still Beats the CPU.” *Wired*, June 2007. <https://www.wired.com/2007/06/ff-humancomp/>.

“Turning Big Data into Actionable Information.” *Mezuro*. Accessed July 8, 2018. <https://www.mezuro.com/turning-big-data-into-actionable-information/>.

[//www.mezuro.com/en/technik/](http://www.mezuro.com/en/technik/).

“Understanding Mobility.” *Mezuro*. Accessed July 8, 2018. <https://www.mezuro.com/en/>.

Vaithianathan, Rhema, Tim Maloney, Emily Putnam-Hornstein, and Nan Jiang. “Children in the Public Benefit System at Risk of Maltreatment: Identification Via Predictive Modeling.” *American Journal of Preventive Medicine* 45, no. 3 (September 2013): 354–59. doi:10.1016/j.amepre.2013.04.022.

Van Hoboken, Joris. “Comment on ‘Democracy Under Siege’: ‘Digital Espionage and Civil Society Resistance’ Presentation by Seda Gürses.” Spui25, July 2018. <http://www.spui25.nl/spui25-en/events/events/2018/07/democracy-under-siege-digital-espionage-and-civil-society-resistance.html>.

Varian, Hal R. “Computer Mediated Transactions.” *American Economic Review* 100, no. 2 (May 2010): 1–10. doi:10.1257/aer.100.2.1.

Von Ahn, Luis, and Will Cathcart. “Teaching Computers to Read: Google Acquires reCAPTCHA.” *Official Google Blog*. Accessed July 22, 2018. <https://googleblog.blogspot.com/2009/09/teaching-computers-to-read-google.html>.

“Wealthfront Investment Methodology White Paper.” *Wealthfront*. Accessed June 24, 2018. <https://research.wealthfront.com/whitepapers/investment-methodology/>.

“What Is Free Software?” *Free Software Foundation*. Accessed August 13, 2018. <https://www.gnu.org/philosophy/free-sw.html>.

“What Is Personal Data?” *Information Commissioner’s Office*, May 2018. <https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/key-definitions/what-is-personal-data/>.

“What Is reCAPTCHA?” *Google Developers*. Accessed July 22, 2018. <https://developers.google.com/recaptcha/>.

“What We Do.” *Data & Society*. Accessed August 13, 2018. <https://datasociety.net/about/>.

“Who Are We?” *Women on Waves*. Accessed July 21, 2018. <http://www.womenonwaves.org/en/page/650/who-are-we>.

Wong, Julia Carrie. “Cambridge Analytica-Linked Academic Spurns Idea Facebook Swayed Election.” *The Guardian*, June 2018. <http://www.theguardian.com/technology/2018/jun/19/aleksandr-kogan-facebook-cambridge-analytica-senate-testimony>.

“YouTube Premium.” *YouTube*. Accessed August 7, 2018. <https://www.youtube.com/pr>

emium.

Zuboff, Shoshana. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30, no. 1 (March 2015): 75–89. doi:[10.1057/jit.2015.5](https://doi.org/10.1057/jit.2015.5).

———. "Google as a Fortune Teller: The Secrets of Surveillance Capitalism." *Frankfurter Allgemeine Zeitung*, March 2016. <http://www.faz.net/1.4103616>.

This thesis is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/) license, making it a part of our global knowledge commons.

