Max Frischknecht
Master Thesis
HS2019/2020
January 9th, 2020

HKB BFH
Master Design
Design Research
Mentoring Ulrike Felsing,
Katherine Hepworth

Predicting Voters:
The Significant Role of Personal Data for Political Communication in Switzerland's Social Networks

1. Abtract

Social networks like Facebook, Twitter and others are becoming increasingly important and enable novel approaches for political communication. Simultaneously, the misuse of personal data is of rising concern for many policymakers worldwide. Firstly, computational automation enables the easy and quick dissemination of opinions, while false social media profiles create the impression of fake publicity. Secondly, personal data collected from social networks offer political actors the possibility to predict the behavior of their potential voters. Methods of psychology are used for target group segmentation and are the basis for persuasive political adverts (Micro-Targeting) like Cambridge Analytica used during the US elections 2016. These developments prompted universities around the globe to launch research projects. However, studies that examine the significance of personal data for political communication in Switzerland are rare to find. Although the elections in October 2019 showed clearly that swiss parties increasingly use personal data for their campaigns. There is little research that sheds light on the involved actors, for example, the agencies and their relationships with the parties.

Therefore, the project addresses the following research questions: **How do Switzerlands parties collect and utilize personal data from potential voters to predict their attitudes, motivations, and behaviors? How does this influence their political communication in terms of design and message? With whom do they collaborate in this process?**

The project produces knowledge about the process of using personal data for political communication on the example of Swiss ballot meetings and analyses the involved actors and their relations. It proposes a combination of different research methods to examine the **outcome (e.g. the advert on Facebook)** as well as **the decision-making process behind the advert.** A web content analysis creates a database of political adverts placed on social networks by Swiss parties. A series of expert interviews illuminate the intentions of the actors (e.g. parties, agencies) who created these ads. The collected data is visualized in three thematic maps that depict the actors, their relations and the political communication in the networks. The project is conducted by an interdisciplinary research team from the fields of political science, data science, and design research. The findings will be disseminated in lectures, peer-reviewed journals, and through an online open-source documentation. The produced knowledge is valuable for state institutions during election observation or social networks preventing digital propaganda.

2.1. State of the Art
Introduction

When compared to other countries, such as the United States, England or Germany, Switzerland has been the subject of few studies that examine the significance of personal data for political communication in social networks. In the following, the state of research on political communication and manipulation in social networks will be introduced. This introduction consists of two stages. The first stage describes the global and local context and concrete methods of political communication and manipulation. In the second stage, the focus is directed onto the process of research itself and research methods are discussed concerning their credibility. This involves discussing data visualization and its importance for working with big data and the contradictory role of using social networks as data sources.

Before introducing the state of the art, a few fundamental terms and concepts should be briefly clarified to avoid misconceptions.

- *Propaganda:* The project understands Propaganda as „persuasive mass communication that filters and frames the issues of the day in a way that strongly favours particular interests; usually those of a government or corporation"(Chandler & Munday 2011). This includes „the intentional manipulation of public opinion through lies, half-truths, and the selective re-telling of history"(Ibid.).
- *Disinformation:* is seen as a practice „involving the dissemination of false information with the deliberate intent to deceive or mislead"(Chandler & Munday 2011). Usually trough mass media.
- *Mass Media:* 1) the technological means of spreading messages and information to a large, widely spread audience, 2) large-scale institutions that produce and disseminate these messages. **Following these definitions, social networks like Facebook and Twitter count as mass media and using them with the intent to influence public opinion in favor of particular interests can be seen as propaganda.**
- *Personal Data:* Defines as „any information that relates to an identified or identifiable living individual[1]". An individual is 'identifiable' if it is distinguishable from other persons. This also includes pieces of information which, when brought together, can lead to the identification of a person. Concerning behavior in a social network, every activity (sharing, commenting, liking, posting, scrolling), can be measured and put into relation with the individual, therefore, becoming personal data. **As a result, when politically contextualized, this data becomes significant for political communication.**
- *Anonymised Data:* Contrary to personal data, this is data where it is impossible to identify individuals. The anonymization must be proven to be irreversible to be anonymous. Meaning it must not be possible to identify the persons again by reengineering the anonymization process.
- *Legal situation:* It is difficult to make general statements about the legal situation. One reason for this is that the respective platforms (e.g. Facebook) are usually not liable in the respective country (e.g. Switzerland). The users of the platforms are usually not aware that they agree to the relatively lax use of their data by accepting the terms and conditions. For example, Cambridge Analytica illegally collected data while the Swiss parties remained within the framework of the local laws. However, it is also the case that software used by some Swiss parties, like Nationbuilder, is again prohibited in France (Püntener 2019).

1 European Commission - European Commission. "What Is Personal Data?" Text. Accessed December 4, 2019. https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en.

Part 1: Political Communication in Social Networks

The Global Context

The use of automatization, algorithms and big data analysis with the intent to influence the political opinion is defined by the Oxford Internet Institute (OII) as „Computational Propaganda" (Bradshaw & Howard 2018). Political and private actors who carry out such propaganda campaigns are termed „Cyber Troops" (Ibid.). The number of affected countries constantly rises and has reached 70 by 2019 (Bradshaw & Howard 2019). Global spending on such campaigns is more than half a billion dollars for the period between 2010 and 2018 (Bradshaw & Howard 2018). The first evidence of the use of Computational Propaganda dates back to 2010 (Bradshaw & Howard 2017), however, conceptually similar political strategies, like the defamation of oppositional candidates, are much older (Grassegger 2019). Today the availability of large amounts of personal data from potential voters enables political actors to create potentially more persuasive and purposeful messages as we will see throughout this chapter. Broader public awareness gained the phenomenon trough a worldwide scandal produced by the consulting agency Cambridge Analytica. The firm harvested personal data from Facebook to create psychologically-manipulative digital advertising for the US presidential elections in 2016 and the Brexit vote (Davies 2015; Krogerus & Grassegger 2016). This incident led to some improvements like the creation of Facebook's Ads Library, where it's possible to look up political advertisements. However, the technological preconditions develop daily and call out for an intense, ongoing commitment. Therefore, institutions around the globe launched research projects to investigate the matter on an ongoing basis. The OII published a series of global reports (Bradshaw & Howard 2017, 2018, 2019) and a set of country-specific reports on Mexico (Glowacki et al. 2018), Sweden (Hedman et al. 2018), the United States (Howard et al. 2017), Brazil (Machado et al. 2018), Germany (Neudert et al. 2017), Ukraine (Zhdanova & Orlova 2019), and many others. The British Information Commissioner's Office (ICO) published reports on the Brexit case (ICO 2018a, 2018b) and New York's University initialized the Social Media and Political Participation Lab.

The Local Context

Despite the described range of research by global institutions, examinations with a focus on the Swiss context are seldom. One exception is the by the Swiss Radio and Television (SRF) published podcast *Hotspot*[2] and short recurring reports in the news show *10vor10*[3] during the votes in October 2019. These reports made clear, that the global developments don't stop at Switzerland's border. The lates votes in October 2019 showed unequivocally, that personal data becomes a driving force in Switzerland's political campaigning (Püntener 2019). Almost every major party used personal data on possible target groups to inform their communication strategy. Be it to decide on which door to knock, which number to call, or, in the case of the FDP, which image-text combination to show to which Facebook user, personal data served as a basis. The reports by the SRF and own research (cf. 2.2) showed, that working with data is a complex procedure. It requires various technical specializations to collect and utilize personal data from potential voters. Consequently, this process also involves different actors, including parties, agencies, data brokers, creators of profiling models and the platforms (e.g. Facebook). However, research examining these actors and their roles in regards to their influence on political communication is yet rare to find. In line with the global trend, spending on digital campaigns is rising among Swiss parties. The mother party of the CVP, for example, invested 260'000 CHF in a digital campaign (Ruch 2019). The success of the Grüne in October 2019, who resigned from working with data because of privacy concerns, also shows that data-based campaigning is not automatically a recipe for success. Experts assume that the trend will continue and the parties will intensify their engagement with data, assumingly becoming more effective in using it (Püntener 2019). Since these developments happened very recently, it does not surprise, that systematic research on Switzerlands situation

2 https://www.srf.ch/play/suche?query=hotspot%20datenspur, accessed January 9th

3 https://www.srf.ch/sendungen/10vor10/wahlkampf-digital, accessed January 9th

is still seldom. The high likeliness of an ongoing, intensified engagement of the parties with personal data calls out for an investigation that enables a broad public discussion around personal data becoming a political commodity.

Now, we start to grasp the global, local, and financial dimensions of the topic. But we didn't look at concrete methods of political communication that profit from personal data and computation. The following sections will introduce some of these methods, however, it is beyond the means of this proposal to create a concluded description of all methods. Rather should two prominent ones, Micro-Targeting, and Social Bots, be described in terms of their state of the art and their role in the Swiss context.

Fake Profiles and Social Bots

One practice of computer-aided propaganda that gained much attention is the creation of false social media profiles, commonly called Bots or Trolls. According to Facebook, such accounts make up 3-4% of all users. Between January and March 2018, Facebook deactivated 538 million of fake accounts. And according to an analysis by cybersecurity firm Imperva, bots accounted for about half of global internet traffic in 2015 (Zweifman 2015).

Of course, not all bots are politically motivated. False or artificial social media profiles, also in the political context, are summarized under the collective term Social Bots. When organized in large quantities, these profiles can give the impression that many, supposedly real citizens, have a certain political opinion. In a political context, they are defined as: „the algorithms that operate over social media, written to learn from and mimic real people so as to manipulate public opinion across a diverse range of social media and device networks." (Woolley & Howard 2016, 4885). Social bots can be divided into three basic types:

1. *Bots:* Purely artificial profiles. All their activities are evoked through computer programs. This includes, for example, the retweeting or sharing of articles from certain politically influenced media portals or the distribution of fake and junk news. Bot activities have been detected in 38 countries. (Bradshaw & Howard 2018).
2. *Trolls:* Profiles operated by humans. They take part in online discussions trying to influence public opinion. A popular example is the „50 Cent Party" which operates on behalf of the Chinese state (King et al. 2017). Trolls are often used to suppress or defame certain opinions through hate speech. Several trolls working together are called a „troll factory". Such movements emerge in a decentralized way over chatrooms or forums. Investigations into the GamerGate affair have shown that there are detailed tutorials available that explain to users how to create false Twitter profiles and effectively combat opposite opinions (Trice & Potts 2018).
3. *Cyborgs:* A combination of artificial and human-based activities. Because of their human component, these bots are especially hard to detect. The OII found evidence of cyborg activities in 9 countries (Bradshaw & Howard 2018).

Depending on the case and the research, further or other bot categorizations are made, including *Impact Bots, Amplifiers, Dampeners, Complainers, Trackers*, or *Service Bots* (Zhdanova & Orlova 2019; Dubois & McKelvey 2019). The detection of bots, meaning the distinction from a profile operated by a real human being, is not always easy to achieve. A fundamental feature is that bots do not produce genuine content of their own, but, for example, simply distribute the content of a news page. The choice of the news page, however, indicates the political motivation behind the bot (Sanovich et al. 2018) (cf. Fig. 1).

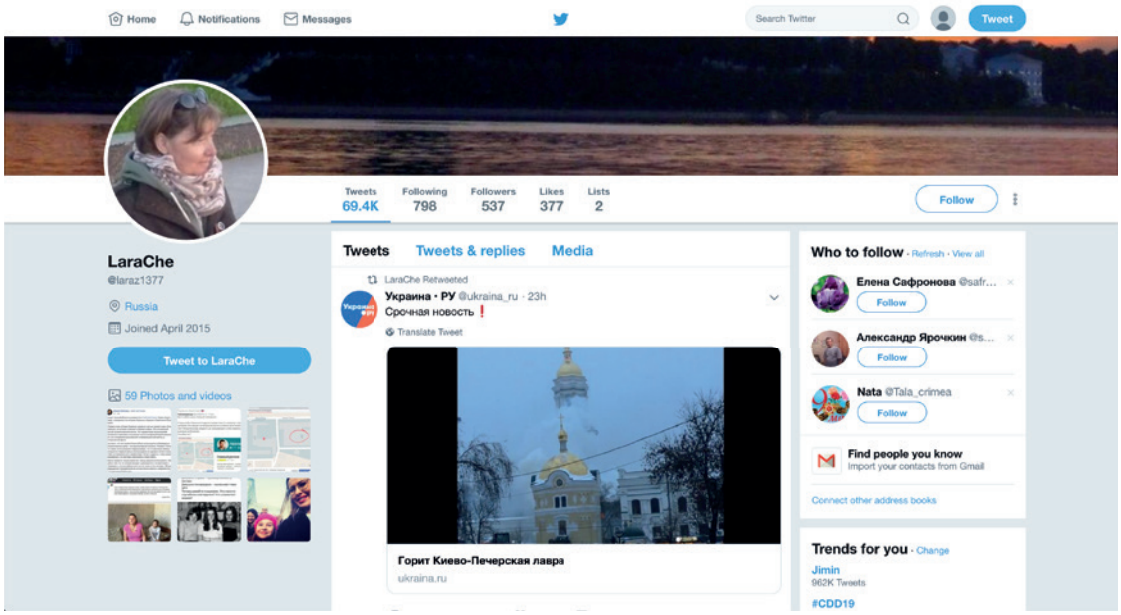


Fig. 1 A bot classified as an Amplifier because it only retweets content from news sources on a high frequency (about 64 tweets per day) (Zhdanova & Orlova 2019)

4 https://botswatch.de

5 https://botometer.iuni.iu.edu/#!/

A series of private and public institutions specialized in the observation and recognition of bots, for example, Botswatch[4] or the Observatory on Social Media[5]. It must be noted that bots are not fundamentally something negative. For example, they are actively used by journalists to collect data or monitor government websites for news (Dubois & McKelvey 2019). I also used a bot to collect tweets from the Swiss parties during the votes in October 2019 (cf 2.2). Till now there are no indications of political bots being active in Switzerland. However, there has not yet been much systematic research on the usage of political bots in Switzerland. It is to note that, even though bots have received a lot of attention, their influence is controversial and many experts believe that bots are less influential than often believed (Reuter 2019). Additionally, since personal data doesn't play a foreseeable role in the usage of bots, the project proposes to turn the focus towards a much more subversive and persuasive way of communication: Micro-Targeting.

Micro-Targeting

Micro-Targeting is an advertising technique that is increasingly used in political communication. It aims to define highly specific target groups based on unique data about that group. In contrast to conventional target group specific advertising, this data goes far beyond socio-demographic characteristics. Not age, income or gender are of primary interest, but **attitudes, motivations, and behavior**. Classic advertisement tries to reach as many people as possible (e.g. women over 30, with children and average income from the Bern region). Micro-Targeting aims for the opposite, reaching the smallest definable groups (e.g. women whose parents are divorced, like punk music and eat vegetarian food). Based on this information the (micro-)targeted advertisement, which can be of various media like a mailing, online advert or phone call, is crafted in a way that **the data suggests is likely to resonate.**

Consequently, a fundamental part of Micro-Targeting is the linking of large amounts of data from potential targets with methods of group segmentation and profiling. For this purpose, the political actors use methods from psychology as the case of Cambridge Analytica and the Swiss votes showed (Krogerus & Grassegger 2016; Püntener 2019). These so-called profiling models help to undertake the data based group segmentation and to make predictions about how likely a target is to respond towards a possible advert. It is assumed that the more personal such data is, the higher the accuracy of segmentation and prediction is. Cambridge Analytica used the Five-Factor-Model (FFM) for this purpose. The FFM is well situated in psychology since the 1980s and defines personality traits along the five basic dimensions of openness, conscientiousness, extraversion, agreeableness, and neuroticism (McCrae & Costa 1997). It's important to understand that, today, social media provides psychology with unprecedented amounts of personal, freely available, data. Because of this, researchers achieved to predict personal attributes like gender, sexuality or political orientation based on Facebook-Likes with an accuracy of 80% to 90% (Kosinski et al. 2013). They state that they can predict a person better than the person's partner based on a data set of 200 likes (Popov 2017). Cambridge Analytica used this exact method to create

Micro-Targeting ads addressing potential voters based on the predictions made from that voter's Facebook-Likes (Krogerus & Grassegger 2016). The agency illegally collected the data on Facebook through a fake play game including a hidden personality test. The Swiss data broker Schober AG, who sold data to the CVP and FDP, uses a similar model to the FFM called „Sinus Milieus" for its analysis (Püntener 2019). The model's typology divides target groups based on „values and views of life" and „social situation". A brief analysis of political advertising on Facebook that I undertook during my studies has shown that, for example, the FDP places ads that suggest the use of profiling methods.

These ads work with a defined set of images and different assigned texts to alter the message. For example, one of the advertisements argues from an ecological perspective and the other from a more economic perspective (cf. Fig. 2–3 and 2.2 *The Network II*). I assume that these different meanings target potential voter groups segmented based on personal data. **We recognize, that using personal data for Micro-Targeting requires various specializations like data collection and analysis and brings different actors into play: the creator of the profiling model (e.g. the Sinus Institute), the data broker (e.g. Schober AG) and the agency designing the adverts.**

Fig 2–3 Political adverts by the FDP placed on Facebook during the votes in October 2019.

We can conclude that computational methods for political communication and manipulation are on the rise worldwide and in Switzerland. While social bots are a large topic, the use and misuse of personal data is of great significance. Research that examines how Swiss parties and other involved actors (platforms, data brokers, creators of profiling models and agencies) **collect and utilize** personal data to predict **attitudes, motivations and behavior** from potential voters is yet rare to find. Not much is known about the **decision-making process** that leads to the design of different adverts who address users individually. However, this is by no means an easy research undertaking, therefore we will look at some research methodologies that analyze communication in social networks regarding their credibility in the next section.

Part 2: The Process of Research

Data Visualisation: A Necessary Research Tool

When examining political communication in social networks, researchers end up with large amounts of data, for example, Tweets in the millions. As a result, many researchers use data visualization as a tool to depict results and uncover patterns in data. As Katherine Hepworth, a specialist in the ethics of visualization notes: „Working with big data shifts the role of data visualization in research. It moves from an optional research output to a necessity for data exploration." (Hepworth 2017, 12). While it is possible to manually examine a few hundred data points, it becomes a „physiological impossibility" if the set consists of millions (Ibid. 12). Data visualizations have proven to be helpful in a wide range of research cases (Tufte 1990). They even have a quite long history

of depicting political power structures and communication patterns (Pfeffer 2017). And they are also used regularly for researching political communication and manipulation (cf. Glowacki et al. 2018; Machado et al. 2018).

However, truthful data visualizations are by no means a self-evident fact. It is rather the case that creating data visualizations must be seen as a complex procedure of engineering, storytelling, and argumentation (Tufte 2001). This process ranges from data collection, revision, analysis, interpretation to final visualization and involves many decisions made by the researcher. These decisions ultimately influence the final argument made by the visualization (Hepworth & Church 2018). Johanna Drucker highlights that „by rendering statistical information into graphical form, a simplicity and legibility is created that hides every aspect of the original interpretative framework on which the statistical data were constructed" (Drucker 2011, sec. 8). The effectiveness of visualization, as Drucker criticizes, has caused humanities researchers to lose the necessary criticism for rigorous scholarly research: „The sheer power of the graphical display of ‚information visualization' […] seems to have produced a momentary blindness among practitioners who would never tolerate such literal assumptions in textual work" (Ibid., sec. 5.). What Drucker describes as an „interpretative framework" refers to the many steps, including data collection, revision, analysis, interpretation, and visualization, to create a chart or map. All the small decisions along the way sum up and formulate the argument of the visualization. Therefore a visualization can't be seen as a simple representation of pre-existing facts as we usually assume. Striving for high standards when creating data visualizations seems like an inescapable necessity, even an ethical duty as some argue (Cairo 2019). This, as I would like to add, holds especially true when visualizing political data in the current climate of distrust, fake news and increasing disbelief towards scientific findings. As we will see soon, the described problems can also be detected in data visualizations depicting results from research on political manipulation in social networks. But before we examine such a visualization of political communication, let's quickly look at some attempts to achieve the described standards and how they could help our research.

One way is to consider the latest guidelines from the field like the by Hepworth (2019) developed Ethical Data Visualization Workflow. Although it is not specifically designed for political content, it can serve as a starting point. Other approaches aim at clearly understanding and contextualizing the data source. **This especially includes acknowledging and depicting the fact that data sets are rarely 100% certain.** This can have different reasons: missing values, sources of low credibility, errors introduced during storage or computational processes, deviations or lack of precision (Griethe & Schumann 2005). Different concepts and typologies try to detect, categorize, and measure uncertainty in data (MacEachren et al. 2005; Thomson et al. 2005). However, these approaches are mainly developed for specialized domains like cartography. They don't emphasize on the topics of social networks in a political context. Nevertheless, these theories help us to detect one problem that visualizations of political communication and manipulation usually have: the uncertainty of the data source. Let us look at this data source and an visualization example in more detail to concretize the mentioned problems.

APIs: Controversial Datasources

To make a data visualization you need data. And when visualizing political communication in social networks most studies collect the data from the network itself. This collection is made possible by accessing the network's Application Programming Interface (API). APIs enable us to communicate with external software (e.g. Twitter) trough code to send and receive data. For example, you conduct a search for the hashtag „wahlen2019" and the Twitter API will return Tweets containing it. You can also receive data about a party's, candidate's or user's profile, and information about the content they share. This data is analyzed by researchers and, for example, visualized in the form of a map of the political affiliation of Twitter profiles (Fig. 4).

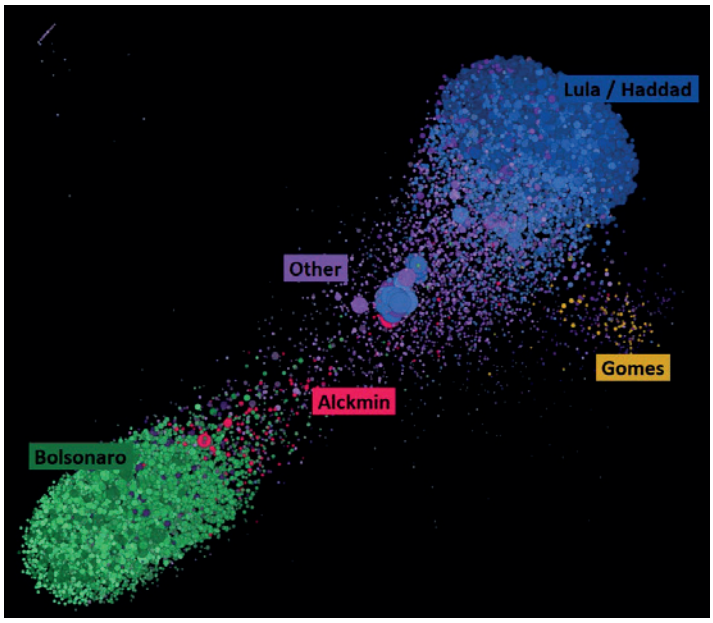


Fig. 4 Twitter profiles affiliated with possible candidates by color and position based on the content these profiles share and like. Each node (dot) represents a profile while its size determines the amount of followers (reach) (Machado et al 2018)

6 https://www.graphika.com/

To draw the above Fig. 4 Machado et al. used the Twitter API as a data source and the commercial visualization software Graphika[6]. However, the representative character of APIs as data sources is controversial. Machado et al. note in their paper: „The platform's precise sampling method is not disclosed, however, Twitter reports that data available through the Streaming API is, at most, 1% of the overall global public traffic on Twitter at any given time" (2018, 3). This means that the API only offers a fraction (1%) of all the possible data from Twitter. **And what the decision-making process is, of a data point becoming part of this 1% or not, remains unknown.** It is, therefore, reasonable to ask, how representative is 1% in regards to the political communication on Twitter? Justifiably there are critical voices from the research community: „A data set made up of any Tweets […] might contain important public discourse about relations […] (but), it is impossible to determine if those Tweets represent all such discourse, what universe of discourse they came from, or the nature of that universe. […] The representativeness of non-random samples, drawn from an unknowable universe, is pure conjecture" (Lacy et al. 2015, 3).

The composition of 1% of the data remains unknown. Often referred to as „Black Box", few people understand what is happening inside the algorithms of big data computation and this also holds true for APIs. The reason for this is that the platforms don't disclose how their algorithms work. Nevertheless, an API is one of the few ways to access a social network. Therefore, API's as data sources can't be avoided completely.

I would like to argue, that we need to shift the focus towards a responsible contextualization of the collected data. This includes **showing these limitations in visualizations.** If you look at Fig. 4 without having read this section you would most likely assume that it represents 100%. All the information you need to comprehensibly „read" the visualization isn't depicted but hidden in the corresponding text. To be able to pay attention to such details in visualizations you need experts from the field of design.

It can be summarised, that data visualizations are a necessity when working with data sets in the millions. The involved steps of data collection, revision, analysis and interpretation shape the argument of the visualization. Data collection on social networks is mostly undertaken through the networks API. However, this API is also a „Black Box" since it's unclear which data can be collected and which not. **This uncertainty is usually not visible in the visualization.** I argue, that in a context of distrust, fake news and increasing disbelief towards scientific findings, **researchers are obliged to show these limitations in their visualizations**, and not „hide" them like it's done in Fig. 4. Otherwise, it's an easy undertaking for critics to doubt research results. To counteract these limitations, research that examines political communication and manipulation in social networks should not focus solely on APIs, but rather expand its sources through conducting expert interviews with the involved actors (parties, agencies, data brokers, etc.). This approach makes it possible to look at the outcome (e.g. the Facebook advert) but also learn about **the decision-making process behind the advert.**

2.2 Current Stage of our Research

2.2 Current State of Our Research

At the current stage, this section summarises research conducted during my MA studies in the field of design. At a later point, this section also includes the state of research from the other members of the research team (cf 2.3.6).

Design Research

During my studies at the University of Arts Bern I examined the political dimension of digitization from four perspectives: 1) the current state of research on Computational Propaganda, Micro-Targeting, and Bots, 2) critical studies on ethical data visualizations, 3) debates on the scientific soundness of digital data sources, for example, APIs, 4) the necessary technological skills to collect, process and visualize digital data, and 5) artistic inquiries from the field of political Media Arts. This acquired knowledge could already, and will further be disseminated trough lectures and workshops:

- "P5js for Beginners", Workshop at FHNW Basel, with Mark Iandovka, February 2019
- "0/1 – Black/White – On/Off – Dead/Alive", Workshop at ETH Zürich, March 2019
- "Copy It Right! On sharing, caring, data and network politics", Workshop, Lecture at Masterstudio Design, FHNW Basel, with Yann Martins, March 2019
- "How to visualize computational propaganda", Lecture at Interaction Design ZHdK Zürich, April 2019
- "Processing", Workshop at Hyperwerk FHNW Basel, with Mark Iandovka, April 2019
- "Fundamentals in digital design practice", Workshop at BA Visual Communication FHNW Basel, February 2020

Scripted Loopholes Series

The five parts of this ongoing series were developed throughout the MA studies and provided a space for technological experimentation. They helped me to grasp the possibilities and limitations of network-related data and to deepen my programming and design skills. The extension of this engagement can be summarised by three main aspects[7]:

1. How and where to collect digital data through the usage of Application Programming Interfaces (APIs). This resulted in experiments with six different APIs, namely the Reuters News API, Twitter API, Wikipedia API, and the by Google developed Geocoding, Maps and Cloud Natural Language API.
2. Learning the necessary technologies to collect, store, revise, analyze and visualize data. This included learning NodeJs, VueJs, D3Js, Git, Heroku and becoming familiar with standard data formats (JSON, CSV).
3. Learning about the limitations of these technologies and more sophisticated data science approaches (Python, Tensor Flow).

Following, the five parts of the series and their relevance for the proposed project should be briefly introduced.

The Search Engine was developed as an assisting tool for my literature and media review. It browses the Reuters database for news articles related to computational propaganda and creates a daily updated news feed. The tool provided first knowledge about how to access an API and how to process and visualize collected data. Through the technical reconstruction of a search engine, it became clear, that search engines like Google can't be regarded as neutral gateways (L'Ecuyer et al. 2018). The search results that Google delivers depends on its construction process. The experiment is a starting point for creating alternative channels of information collected during the proposed literature and media review (cf. 2.3.4). Most leading news and academic outlets (e.g. JSTOR) maintain an API to bypass default search engines.

7 More detailed informations can be found in the documentation under „Technology Excursion"

2.2 Current Stage of our Research

The Advertiser analyses the Facebook Ads Interests, a list that Facebook creates for every profile. This list is partly based on direct user behavior (liking, commenting), and partly predicted by a „Black Box" algorithm. The tool calculates how many entries Facebook created autonomously by comparing the list to my likes:

- Advertisement interests in total: 218
- Liked by me: 32
- Defined by Facebook: 186

Facebook's prediction of my interests was sometimes true but mostly completely wrong. It supported the hypothesis that algorithms are good at collecting but bad at contextualizing. Nevertheless, I still believe it's possible to make assumptions about my attitude, motivations, and behavior from this list. A relatively large amount of terms in the list where unknown to me, therefore, the tool includes an automated Wikipedia summary for every term. The tool helped to deepen the knowledge about social network-related data and how to analyze it.

The Messenger is inspired by the uncovering of Facebook selling private chat records to companies like Netflix and Amazon (Newton 2018). Such chat records can be of great length and I assume that the companies use algorithms to extract useful information. I was curious how much information an algorithm could derive from a chat between a friend of mine and me. Inspired by the discovery of my ad interests in *The Advertiser*, I chose to run an entity analysis on my chat with a Google-developed machine learning algorithm. The algorithm identifies dates, persons, contact information, organizations, locations, events, products, and media types in unstructured text. Surprisingly the produced results consisted of almost 90% noise (wrong predictions). Nevertheless, it was possible to assume the personal address, profession, and country of origin of my friend from the results. The tool helped to create knowledge on machine learning-based APIs and text analysis. It could be potentially helpful for data analysis of expert interview transcripts (cf. 2.3.4).

The Network is an ongoing project to visualize political activity on Twitter during the Swiss votes. The tool uses the Twitter API and supports the thesis that APIs only provide a small insight into a network (cf. 2.1 APIs). Between the 17th and 20th of October, I collected around 12'000 Tweets with the help of a data collection script. Simultaneously this provided me with the necessary skills to program a Twitter-Bot. The search query consisted of political profiles from parties and candidates and political hashtags like „#jetztFDPwählen" or „#darumstarkeSP". The collected Tweets were pruned and geographically mapped across Switzerland. The project has shown that more complex undertakings usually require the use of different APIs to collect, analyze and visualize data. For example, was the Twitter API used to collect data, while two APIs by Google helped to define the Tweet's geographic origin and placing it on the map of Switzerland. A more detailed description of the process can be found in the documentation (cf. SL4 The Network Logs 1-6).

The Network II is an ongoing investigation of the Facebook Ads Library. Research by the SRF has shown, that the FDP obtained data from Schober AG via the opinion research institute GSF Bern (Püntener 2019). In the advertising library of Facebook, various Swiss election advertising can be found. A series of adverts placed by the FDP show several different design alternatives making usage of versatile image-text combinations (Fig. 5-8). I assume that the different arguments of the advertisements aimed at different (micro-) target groups. The by Schober AG used „Sinus Model" (cf. 2.1 *Micro-Targeting*) could have served for the segmentation of the groups. These findings serve as a starting point for a preliminary study that I will submit to the BFH Call for Proposals 2021 (cf 2.5).

2.2 Current Stage of our Research

Fig. 5–8 Political adverts by the FDP placed on Facebook during the votes in October 2019.

Relevance for the research project

The engagement with the Scripted Loopholes series supports the hypothesis that the collection, analysis, and visualization of data is a complex procedure calling for an interdisciplinary team of specialists. Supporting the thoughts of Drucker (2011), Hepworth & Church (2018), collecting data and creating truthful data visualizations is by no means a self-evident process. **The more complex an investigation becomes, the more technologies are required for processing.** Each of these technologies, in turn, has the potential to have an impact on the nature of the data and the argument of the visualization.

The first four parts of the series can be visited together with the documentation under: *https://mafiles.maxfrischknecht.ch*

2.3 Detailed Research Plan

2.3.1 Assumptions & Hypotheses

State of the art and the own research findings lead me to the following hypotheses: Personal data becomes significant for the political communication of the Swiss parties while the involved actors and their methods of collecting and utilizing personal data are intransparent. This lack of transparency poses the danger, that a switch from contemporary data-informed campaigning into computational propaganda goes unnoticed by the public. The line between legally using the available technological possibilities for political purposes and the unethical, manipulative misuse of personal data is thin and needs to be defined. The role of personal data differs very much based on the involved actors and the specifics of data collection and utilization, therefore, knowledge applicable to the swiss context is limited. To improve transparency it is necessary to acquire the following knowledge trough rigorous on-site research:

- If we want to understand the role of personal data for political campaigns we can't only look at the outcome (e.g. an advert on Facebook), we also need to investigate the circumstances and the decision-making processes that created this advert.
- Many actors with different roles are involved in this process and it's necessary to identify them. At this stage, the following actors have been defined:
  - *Users* of social networks who spread personal data
  - Platforms who collect, analyze and sell the data of their users (e.g. Facebook)
  - *Data brokers* who aggregate, analyze and sell data (e.g. Schober AG)
  - *Creators of profiling models* that help to undertake the data based group segmentation (e.g. Sinus Institute)
  - *Communication agencies and advisors* who conceptualize and design the communication (e.g. Farner, Enigma)
  - *The party leadership* that develops and implements the overall strategy
  - *Individual candidates* of the parties who develop independent communication strategies for themselves
- These actors have different motivations and expectations for working with data. While the creator of a **profiling model** might be led by scientific interest, the data broker could be guided by economic and the party by political interests. This influences how they process the data because they expect different results from working with it.
- It's important to understand how personal data enables the actors to make predictions about a person's **attitude, motivation, and behavior**.
- The ability to make such predictions changes the design process of the adverts. Without this data the agencies have to guess what resonates with the voters, now they can claim that they know beforehand.
- It's important to understand the transformative processes and the different characteristics of the data during the process. While it could be the case, that personal data used to create a profiling model (e.g. the Sinus Model) is anonymized, therefore not personal anymore, the data used by a data broker to feed the model, however, is personal.
- It's important to understand what characterizes the relationships these actors have with each other, are they commercially, politically or otherwise? For example, it would be critical if an institute with a public mandate would share its findings with a party.
- Understanding the intentions and the processes behind the outcome (e.g. the Facebook advert) are much more valuable than just analyzing the different outcomes itself. It leads to a more profound understanding of the meaning and power of personal data for political communication by looking at the decision-making process. While analyzing only the outcome leaves the core of the matter blurry and in the worst-case contributes to polarising opinions. **An understanding of the background, therefore, is also more likely to contribute to a differentiated debate.**

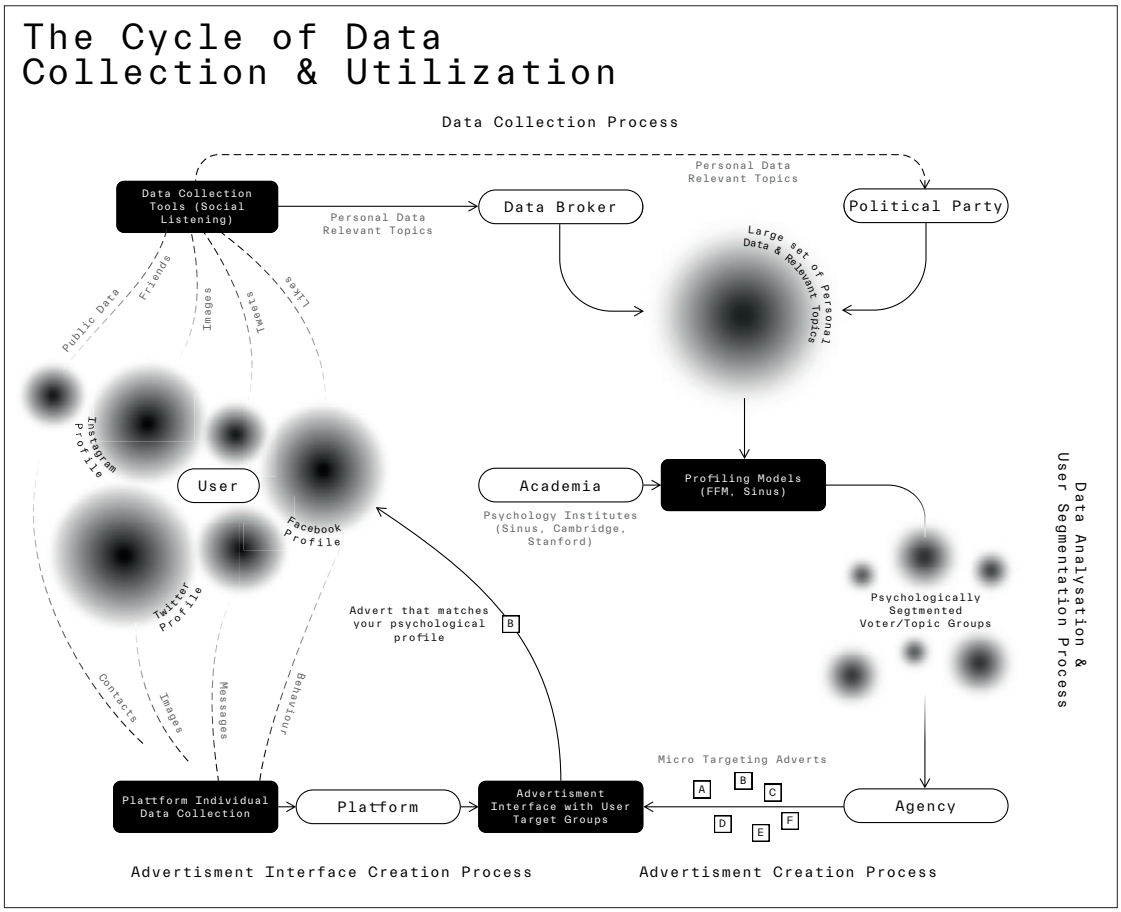


Fig. 9 Actors Map

These assumptions are the basis for the Actors Map (Fig. 9). It shows the presently identified actors, their relationships and the processes of collection, analysis, interpretation and communication. This visualisation serves the project as a starting point (cf 2.3.4 Data Visualization).

2.3.2 Research Question

How do Switzerlands parties collect and utilize personal data from potential voters to predict their attitudes, motivations, and behaviors? How does this influence their political communication in terms of design and message? With whom do they collaborate in this process?

To answer the overall questions the project formulates a set of subquestions:

- What actors are involved in the process of, and how do they collect, revise, analyze, and interpret personal data for political communication?
- What are the expectations and motivations of the actors to work with personal data?
- How is the character of the data at the stages of collection, revision, analysis, and interpretation, and how does it transform from one stage to the other? (Is it always personal? When does it become personal?)
- What conclusions are drawn from personal data and how does it influence their political communication in terms of design and message?
- What constitutes the relationship the actors have with each other, are these commercially, politically or otherwise?

2.3.3 Concrete Goals

The proposed project systematically explores the role of personal data for political campaigning in Switzerland. It produces detailed and genuine knowledge about the processes of data **collection, revision, analysis, and interpretation** of the example of Swiss ballot meetings. Furthermore, the project produces knowledge about how personal data enables the parties to predict **attitudes, motivations, and behaviors** from potential voters and how this **shapes the political communication in terms of design and message** (e.g. a Facebook advert). The project clarifies **which actors** are involved in this process, what their **expectations and motivations** are to work with personal data in a political context and **what characterizes the**

8 e.g. Oxford Internet Institute, Political Data Science University of Munich, Social Science Research Council
9 e.g. ICO
10 e.g. Algorithmwatch, Mozilla Foundation, Digitale Gesellschaft, Chaos Computer Club, Ranking Digital Rights
11 e.g. Netzpolitk, Republik, Department of Data Journalism at the SRF where I already established a contact.
12 e.g. Novalityca, Datahouse, Schober AG, GSF Bern, Nationbuilder, Sinus Institute, Farner Consulting, Enigma, Furrerhugi, Facebook/Twitter for Business

relationship between the different actors. In this way, the project contributes to a concrete understanding of the political significance of personal data in Switzerland and places the Swiss situation in a global context. Besides, the project develops new subject-specific knowledge on **how to classify APIs as data sources** and develops data visualizations that take the fact into account that **data is constructed and not the representation of a priori existing facts.** Thus it contributes to discourses in the fields of Political Science, Data Science, Design Research and the Digital Humanities. The research team will disseminate the findings in lectures at international congresses, in peer-reviewed journals, in the form of online open-source documentation and the final publication.

2.3.4 Overview of Methods

Literature & Media Review (WP1)

The literature and media review helps to develop theoretical foundations and to refine hypotheses. It provides technological knowledge (facts and dates) by reviewing literature and key documents. It addresses the following issues:

1. Understanding the global and local context of the role of personal data in political communication. This involves reviewing academic and scholarly studies[8], government reports[9], research conducted by civil society organizations[10] and investigative journalistic reports[11]. It helps the project to assess the significance of personal data and to understand the methods of political communication that utilize personal data (e.g. Micro-Targeting). It guarantees the last topicality of the research, and allows the embedding of the findings in a global context, thereby contributing to the ongoing academic discussion in the field.
2. Understanding the current state of possibilities to retrieve or buy personal data in Switzerland and available offers for accompanying services like analysis, utilization, communication, and design. Trough reviewing service offerings and documents from platforms, data brokers, creators of profiling models and communication agencies[12] the *Actors-Map* (cf. page 14) will be refined. This deepens the hypotheses on the involved actors, their roles, expectations, and motivations to work with personal data.

The *Actors Map* and hypotheses developed upon the review lay the basis for the expert interviews. It enables the project to put the focus of the following expert interviews on process and interpretation knowledge and less on technical knowledge, which, according to Bogner et al. (2014, 18) is consistent with the methodological strengths of the method. However, review and interviews are not understood as linear and consecutive, but rather as mutually influencing and iterative processes.

Expert Interviews (WP1 / WP2 / WP3)

22 expert interviews will be conducted with the following representatives:

- 3 different data brokers (e.g. Schober, Novalytica, Datahouse)
- 3 creators of profiling models (e.g. the Sinus Institute, TBD, TBD)
- 3 different agencies (e.g. Farner, Enigma, Furrerhugi)
- 1 leading representative of each party (SVP, FDP, CVP, GLP, SP, Grüne)
- 1 candidate that works with data in his campaigns (SVP, FDP, CVP, GLP, SP, Grüne)
- 1 representative of each platform (Facebook, Twitter)

The expert interviews provide knowledge about the **decision-making process behind political communication** (e.g. a Facebook advert). It helps to understand the role of the different actors and how they **collect and utilize personal data** to **predict attitudes, motivations, and behaviors from potential voters.** It also provides insight into the actor's expectations and motivations to work with personal data and their relationships with each other. These revelations and roles will be visualized in the *Actors Map*.

2.3 Detailed Research Plan

The expert interview is a method of qualitative social research whereby experts are defined as persons who possess knowledge that is „particularly effective in practice and thus becomes a guide for orientation and action for other actors" (Bogner et al. 2014, 14). Three different types of expert interviews will be conducted to open up different levels of knowledge:

- Explorative Expert Interviews (WP1) serve as orientation in the field and for scrutinizing hypotheses.
- Systematizing Expert Interviews (WP2) try to collect the technological and process knowledge comprehensively and analytically. Concretely they will reveal how the actors collect and utilize personal data for political communication to predict attitudes, motivations, and behaviors from potential voters.
- Theorizing Expert Interviews (WP3) aim at the subjective dimension (implicit decision, perceptual patterns, world views, etc.). They will reveal the actor's expectations and motivations to work with personal data and the relationship the actors have with each other (commercially, politically or otherwise).

Web Content Analysis (WP1 / WP2 / WP3)

The Web Content Analysis (Herring 2010) provides the project with real-world data collected through Facebook's Ads Library API and the Twitter API. The Data sets consist of adverts, announcements and messages from political parties placed in these networks and related metadata like user interaction, links, the period of publication, number of possible views, estimated gender of recipient and locality of the display. It enables the project to analyze 1) the possible target groups of the communication, and 2) design and message of the communication. The collected data will be depicted and analyzed with the help of data visualizations. For each network, an individual visualization will be created (cf. next section, *Facebook* and *Twitter Map*). The findings can then be compared against the findings from the expert interviews who examine the decision-making process behind the communication. Thus, both methods complement each other in the effort of developing a comprehensive understanding of how personal data influences the political communication of Swiss parties.

WebCA considers the information to be contained inside various types of media fragments like links, exchanges or features. It is a methodological plural paradigm and draws from methods like Computer-Mediated Discourse Analysis, which has been applied to analyze forms of dialogue (chat rooms, text messaging) (Herring 2010., 238) or Social Network Analysis, which has been used to analyze political linking practices (Foot et al. 2006). It, therefore, allows for the analysis of various types of information related to social network data (e.g. text, image, metadata).

To define the corpus of research a Web Sphere (Foot & Schneider 2002) is created for every major party (SVP, FDP, CVP, GLP, SP, Grüne), consisting of the party's presence on Facebook and Twitter. It is defined as a „hyperlinked set of dynamically defined digital resources spanning multiple Web sites relevant to a central theme or object" (Ibid. 225) (cf. Fig. 10)

The project limits the data collection to Facebook and Twitter. Other platforms like Instagram are excluded. Facebook's and Twitter's global reach make them still the most dominant platforms for political communication (Bradshaw & Howard 2019). The project can serve as a basis to explore other platforms in future research.

Data Visualization (WP1 / WP2 / WP3)

The project develops three thematic visualizations (maps) throughout the research process to gain new insights regarding actor roles and political communication in networks:

1. The *Actors Map* (cf. Fig. 9, page 14) visualizes the involved actors, their roles and relations, including the user, data brokers, parties,

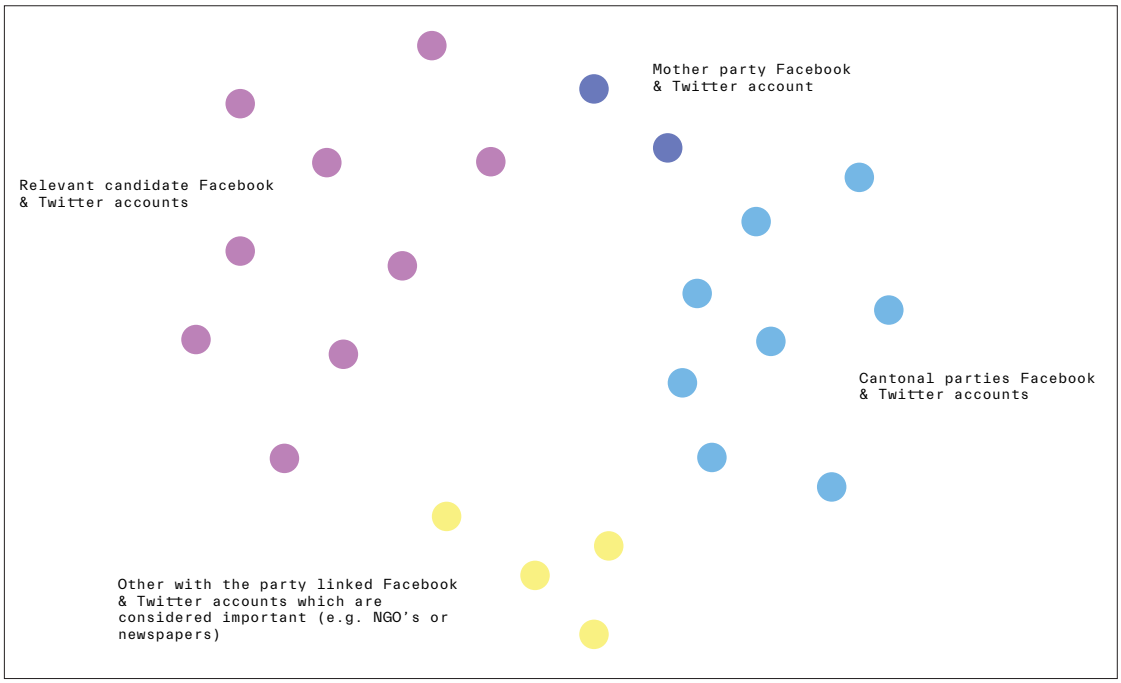


Fig. 10 Web Sphere consisting of party affiliated Facebook/Twitter accounts.

agencies, and the platforms. It visualizes how personal data is collected, processed and utilized for political communication. The map focuses on the decision making processes behind the adverts trough depicting key findings from the review and the expert interviews.

2. & 3. *The Facebook and Twitter Maps* visualize the political communication of the parties inside the respective networks. The maps are based on the data collected through the WebCA (cf. 2.3.4). It enables the project to analyze the design and message of the communication and to put it into relation with corresponding metadata about the target groups (e.g. gender, locality, time). It furthermore enables the project to analyze how the strategies of the parties differ in regards to frequency, range of coverage, medium (image, text, video), message, financial expenses and how the users in the network interact with it.

Every visualization is divided into three parts commonly used in political network visualizations (Pfeffer 2017):

1. *Substance:* Preparing the content/data, i.e. specific information on the actors, data collection/utilization processes and political communication.
2. *Design:* The process of transferring the substance (content/data) into graphic elements. Here the visual means (e.g. color, size), as well as interactivity, is defined (e.g. show/hide layers of information, zoom into different dimensions of the data).
3. *Algorithms:* Development of the computer-based support necessary for the processing and presentation of the data. They help, for example, to reduce complexity by combining data or to position elements spatially based on defined laws. To improve the visualization gradually the project proposes a rapid prototyping approach conducting iterative revision cycles of assessing needs, setting objectives, prototyping and utilizing. (Baek et al. 2004). The Ethical Data Visualization Workflow (Hepworth 2019) will serve as a guideline to ensure that the visualizations meet ethical standards. Other principles discussed, such as visualizing uncertainty (Griethe & Schumann 2005), are also taken into account.

The different teams will develop the maps collaboratively (cf 2.3.7). The regular exchange within the team ensures a development based on the team member's needs. At the end of the project, the visualizations will be published in the open-source documentation.

2.3.5 Work Packages (WP) & Milestones (M)

WP 1 Exploration Phase		
A) Theory	B) Data	C) Design
1.1 Reviewing global/local context on data-informed political communication 1.2 Refining party/profiling actors	1.1 Reviewing data brokers and platforms 1.2 Refining data broker/ platform actors	1.1 Reviewing local agencies working with political data 1.2 Refining agency actors
• Combining actors, describing characteristics and relationships • Define objectives of exploratory interviews, Actors Map, Facebook data collection, Facebook-Map		
1.3 Conducting explorative expert interviews 1.4 Analyzing actors based on interviews 1.5 Mapping actors (in collaboration with C)	1.3 Setup Facebook API access, review current best practice, defining Web Spheres 1.4 Collecting/pruning Facebook data set 1.5 Processing Facebook data set (in collaboration with C)	1.3 Conceptualization of visualizations (maps) and team-user scenarios based on the defined objectives 1.4 Developing Actors Map (in collaboration with A) 1.5 Developing Facebook Map (in collaboration with B)
M1: Q4 2021 Explorative expert interviews conducted M2: Q4 2021 Facebook API accessed, data set collected, pruned, and analyzed M3: Q4 2021 Visualizations conceptualized, Actors Map and Facebook-Map prototyped		
• Based on conducted interviews, Actors Map, Facebook-Map: define objectives of systematizing interviews, Twitter data collection, Twitter Map		

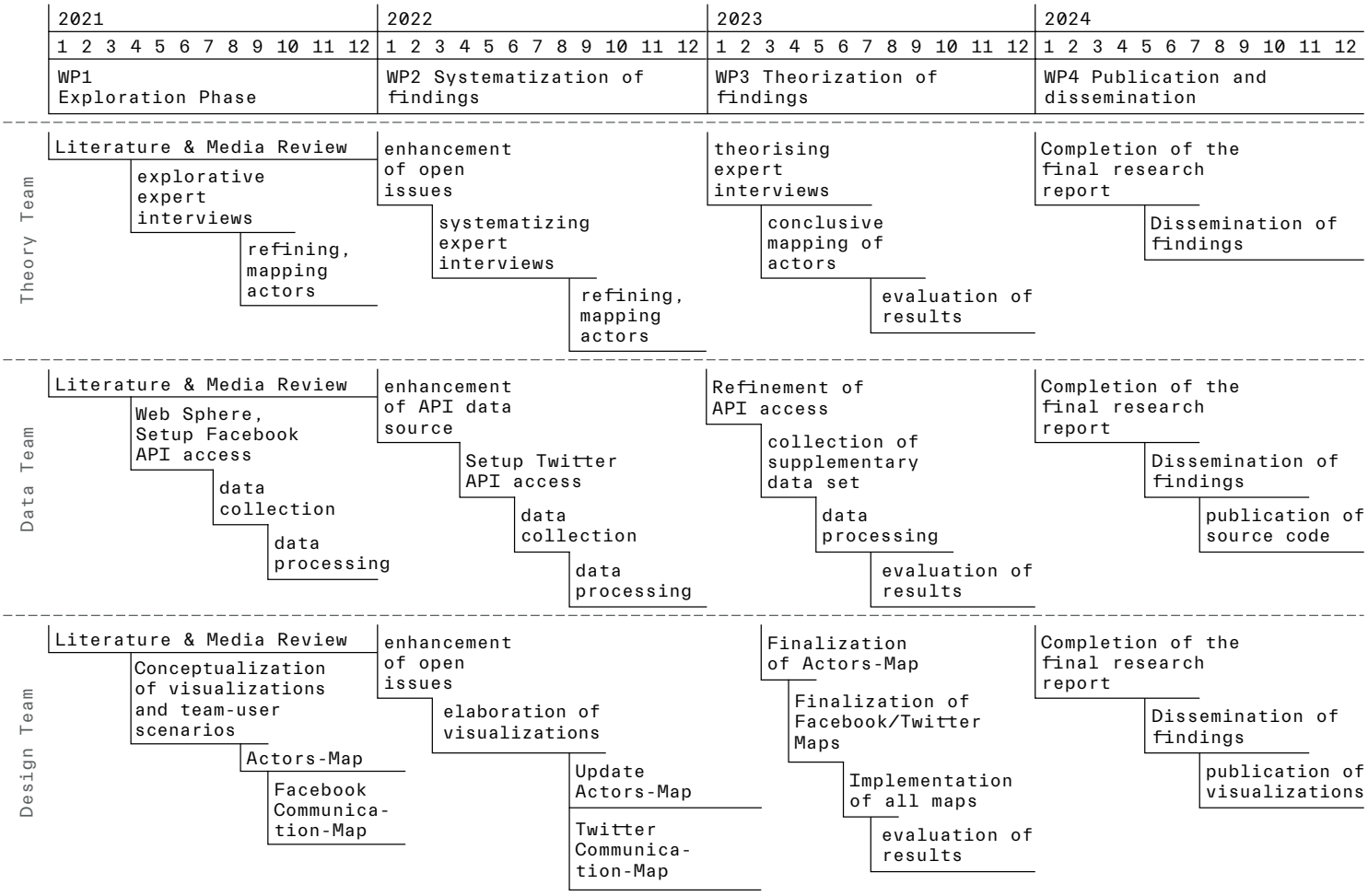
WP2 Systematization of findings		
A) Theory	B) Data	C) Design
2.1 Identification and enhancement of open issues in regards to involved actors 2.2 Conducting systematizing expert interviews 2.3 Refining and describing actors 2.4 Mapping actors (in collaboration with C)	2.1 Identification and enhancement of open issues in regards to API data source 2.2 Setup Twitter API access 2.3 Collecting/pruning Twitter data set 2.4 Processing Twitter data set (in collaboration with C)	2.1 Identification and description of open issues regarding the maps 2.2 Elaboration of maps 2.3 Update and refinement of Actors Map (in collaboration with A) 2.4 Developing Twitter Map (in collaboration with B)
M4: Q4 2022 Systematizing expert interviews conducted M5: Q4 2022 Twitter API accesses, data set collected, pruned, and analyzed M6: Q4 2022 Actors Map updated and refined, Twitter Map prototyped		
• Based on conducted systematizing interviews, updated Actors and Twitter Map: define objectives of theorising interviews, supplementary data set, visualization implementation		

2.3 Detailed Research Plan

WP3 Theorization of findings		
A) Theory	B) Data	C) Design
3.1 Conducting theorizing expert interviews 3.2 Conclusive mapping of actors (in collaboration with C) 3.3 Evaluation and formulation of the results of the expert interviews	3.1 Refinement of API access 3.2 Collection/pruning of supplementary Facebook/Twitter data set 3.3 Processing supplementary data set (in collaboration with C) 3.4 Evaluation and formulation of data collection process and APIs as data sources	3.1 Finalization of Actors Map (in collaboration with A) 3.2 Finalization of Facebook/Twitter Map (in collaboration with B) 3.3 Implementation of all maps, preparation for publication 3.4 Evaluation and formulation of the maps and their functionality
M7: Q4 2023 Interview, data collection process, and maps evaluated and results formulated, dissertations completed		

WP4 Publication and dissemination		
A) Theory	B) Data	C) Design
4.1 Completion of the final research report (SNFS) 4.2 Dissemination of findings in lectures at two international conferences (such as IEEEVis and ACM SIGGRAPH), in two peer-reviewed journals (such as Communication Review Quarterly and Fibreculture), in the form of online open-source documentation and the final publication 4.3 publication of source code in an online-code repository		
M8: Q4 2024 final research report and conference concluded, dissemination of findings initiated		

2.3.5 Timetable



2.3.6 Research Team Organization

The research group unites 8 researchers from different fields and is divided into three teams each consisting of an experienced team leader with a PhD in the respective Field and a PhD candidate. One additional member is taking care of direction and overall concept, another one of planning and coordination.

A) Theory Team

The theory team is responsible for the contextualization, organization and conduction of the expert interviews. The members of the team are **political scientists** with a specialization in political communication and propaganda theories. The team members have a profound understanding of Switzerland's political landscape.

B) Data Team

The data team is responsible for accessing, collecting, pruning, and storing the network data (WebCA). The members of the team are **data scientists** familiar with the most used languages for data analysis (R, Python, SQL, NodeJs, others).

C) Design Team

The design team is responsible for the conduction of the visual prototyping (*Actors, Facebook*, and *Twitter Map*). The members of the team are **designers with specializations in data visualizations, UX, and design research**. They are experts in data visualization tools like R and D3js and familiar with programming.

Interdisciplinarity

During the literature and media review the teams collaboratively define the involved actors. A and C collaboratively develop the *Actors Map* based on the conducted expert interviews. B and C collaboratively develop the *Facebook/Twitter Map* based on the conducted WebCA. Every two to three months of interdisciplinary work meetings with all teams will take place, ensuring discussion about the process and the state of the research work.

2.3.7 Risks

The report by the SRF and personal exchange with journalists showed that agencies and data brokers are potentially open for interviews. However, it is important to acknowledge that some interview partners could not be willing to share their knowledge or have no substantial knowledge about the processes of data collection and utilization. The platforms most likely won't share deeper insights into their data processes, and neither are all political actors willing to give an interview. But, it's also important to note that some parties (e.g. Grüne) are critical towards using personal data. I, therefore, assume that they are interested in the findings of the proposed research and open for the exchange of information. General openness is expected from the creator of the profiling models who come from an academic context (e.g. the Sinus Institute). Either way, enough time is calculated to expand the network of possible interview partners. Additionally, the research does produce valuable insights into the political communication of Swiss parties in social networks through its accompanying research methods. The project no only relies on expert interviews as a data source but also conducts a web content analysis that will provide knowledge about target groups, design, and message of political adverts. A one-year preliminary study opens up important contacts to actors even before the start of the project (cf. 2.5).

2.3.7 Budget
All Prices are in Swiss Francs.

Material costs		
Travel expenses interviews	22 à 100.—	2’200.—
Scientific articles in peer-reviewed journals	2 à 3'000.—	6’000.—
Translations of scientific articles (1 MS=50.–)	2 à 2'200.—	4’400.—
Conferences (travel costs, each team once a year)	per person 500.—	12’000.—
Travel expenses for meetings	per year 2'000.—	8’000.—
Transcription/analysis software for interviews (e.g. Nvivo 12 Plus)		1'200.—
Total material costs		33’800.—

Final publication	
Publishing services	5’000.—
Translation	8’000.—
Copy editing	3’000.—
Proofreading	1’000.—
Prepress	1’000.—
Digitization, Open Access publication	2’000.—
Image editing	1’000.—
Image rights	1’500.—
Print	10’000.—
Layout	15’000.—
Total final publication	47’500.—

2.3 Detailed Research Plan

Personnel costs			
Direction	40%	per year 40’000.—	160’000.—
Planning and coordination	40%	40’000.—	160’000.—
Team A leader	80%	80’000.—	320’000.—
Team A PhD candidate	100%	47'040.— 48'540.— 50'040.— 50'040.—	195’660.—
Team B leader	80%	80’000.—	320’000.—
Team B PhD candidate	100%	47'040.— 48'540.— 50'040.— 50'040.—	195’660.—
Team C leader	80%	80’000.—	320’000.—
Team C PhD candidate	100%	47'040.— 48'540.— 50'040.— 50'040.—	195’660.—
Total personnel costs			1'866’980.—

Total costs	
Material costs	33’800.—
Final publication	47’500.—
Personnel costs	1'866’980.—
Total project costs	**1'948’280.—**

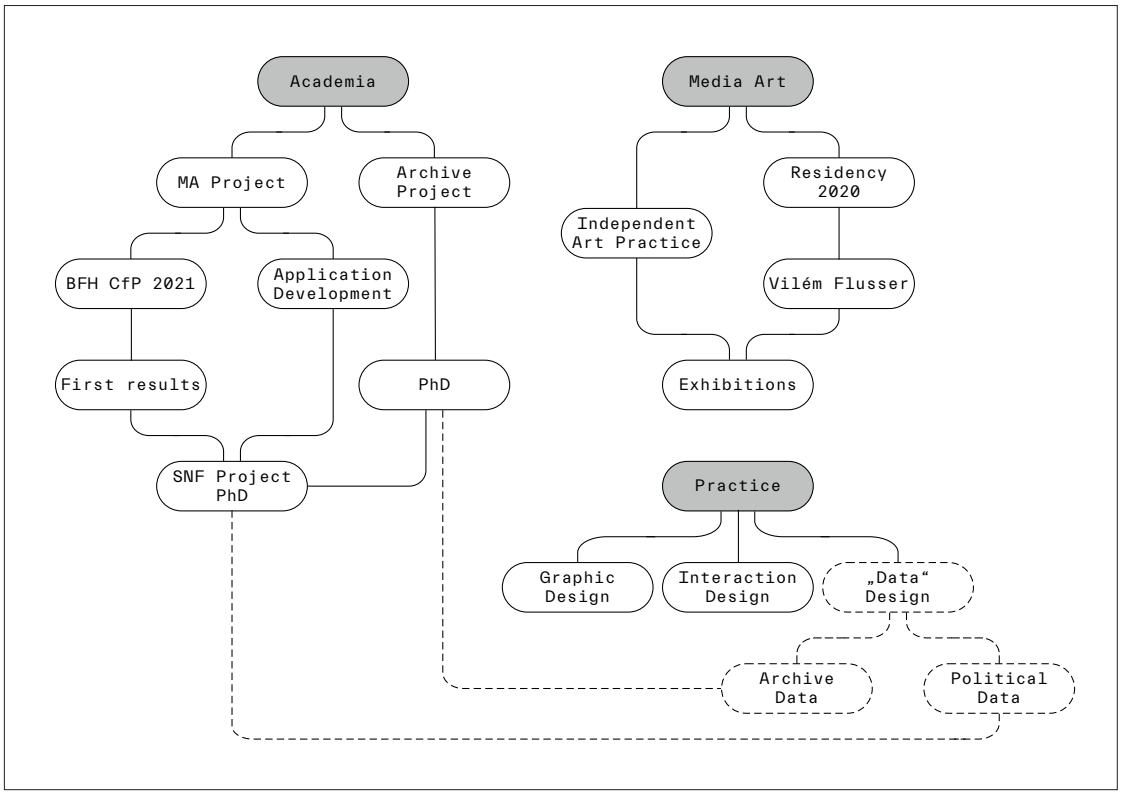
2.4 Relevance & Impact

Scientific Relevance

The project provides crucial insights into the significance of personal data for political communication in Switzerland's social networks. The scientific community still lacks detailed knowledge about **how Switzerland's parties collect and utilize personal data from potential voters to predict their attitudes, motivations, and behaviors.** Other studies, e.g. by the OII (cf Bradshaw & Howard 2019), focus primarily on understanding the effect and extent of digital propaganda and less on the **influence of personal data on political communication** in terms of design and message. Furthermore, **the involved actors, their relations to each other and their motivations** to work with personal data are treated marginally at best. The examination and visualization of using personal data for political communication is for the first time the subject of an in-depth investigation of the example of Swiss ballot meetings. This study not only looks at the macro-level of political communication through an extensive analysis of political advertisements in social networks but also at the micro-level through examining the **decision-making process behind the data-informed campaigns**. This knowledge is considered essential for a differentiated debate about the use of personal data for political means. The study contributes to discourses in the fields of Political Science, Data Science (specifically in regards to APIs as data sources), Design Research (specifically in regards to data visualizations) and the Digital Humanities. The results are to be disseminated through lectures, articles, the online open-source documentation and final publication.

Broader Impact

As a result of the recent scandals around Cambridge Analytica and given the much-cited 'over-complexity' of digital technology, there is an increasing need for improved media literacy and critical ability. The project nourishes this discussion by providing a comprehensive understanding of personal data being used in political processes and thus makes a contribution to society beyond the realm of science. The project raises the question of who has the right to collect and utilize personal data with which intensions. In this way, the project refers to one of the most important contemporary issues, **namely if we as a society want personal data to be a commodity** and how this interferes with basic rights like the protection of privacy. Against the background of current practices and reform attempts in dealing with digital media (e.g. decentralized networks, peer-2-peer communication), the research results promote insights in a field that has so far been treated highly asymmetrically. The complexity of the matter, only allow small specialized groups of programmers and global tech companies an understanding. Due to the rapid technological development, there is hardly any room for moments of reflection in which society defines the social and political requirements social networks have to meet. Various social actors can benefit from an understanding of the processes of political communication and manipulation: State institutions concerning election observation, initiatives for political transparency, social networks preventing political manipulation or educational institutions in the respective fields.



2.5 Next Steps

The following section describes the further procedure after the successful completion of the Master's degree. The engagement with the political dimension of the digital is reflected in three fields of activity, namely academia, practice and media art.

Academia

The project is submitted in smaller scale to the BFH Call for Proposals 2021 in collaboration with Dr. Ulrich Fiedler (BFH TI) who assist as technological expert. The preliminary study analyses Facebook adverts placed by the FDP in regards to their target group, design and message. Simultaneously the research plan is developed further to be submitted to the SNF. The SNF is seen as the right funding partner because the project conducts basic research.
Sinergia, NRP or NCCR are considered appropriate programs to submit the proposal because of it's interdisciplinarity, actuality and relation to Switzerland. Additionally, I became part of the project *Participatory Knowledge Practices in Analog and Digital Image Archives* which has been submitted to the SNF. It offers the opportunity to obtain a PhD in Digital Humanities (University of Bern). My primary supervisor would be Prof Dr. T. Hodel while the three co-applicants of the project (W. Leimgruber, P. Fornaro, U. Felsing) would share the role of the second supervisor. In this context, U. Felsing and I develop a joint article around the topic of data visualization. Additionally, I'm discussing a joint article with my other mentor Katherine Hepworth.

Practice

I will include the diverse valuable experiences of the MA Design in my practice (Studio Début Début) in the coming year. Concretely this means a revision of the business plan that reflects the increasingly scientific way of working and the clientele. Additionally, this includes the development of a new service offer dealing with the processing and visualization of data, specifically in the two contexts of archives and politics.

Media Art

Influenced by political artists like Wachter/Jud or !Mediengruppe Bitnik, I discovered art as a potential space for reflection and mediation of my research findings. The Scripted Loopholes series is has been exhibited at the gallery for contemporary art E-Werk in Freiburg during Regionale19. The series was also proposed as a project at the Vilém Flusser Residency 2020, which would take place between May to June 2020 at the University of the Arts Berlin. The residency would provide the opportunity to develop the series further.

3. Documentation

A complete, annotated documentation of my master studies can be found under: www.mafiles.maxfrischknecht.ch

4. Declaration

I hereby declare that I wrote the present Master thesis myself and that I did not use any other means than those listed, that I did not claim authorship of a text and that I did not make unauthorized use of scientific texts or data.

Max Frischknecht, January 9th, 2020

5. Bibliography

Baek, Eun-Ok, Kursat Cagiltay, Elizabeth Boling, and Theodore Frick. "User-Centered Design and Development." In Handbook of Research on Educational Communications and Technology, edited by David Jonassen and Marcy Driscoll, 13. New York, NY, USA: Routledge, 2004.

Bogner, Alexander, Beate Littig, and Wolfgang Menz. Interviews mit Experten: eine praxisorientierte Einführung. Qualitative Sozialforschung. Wiesbaden: Springer VS, 2014.

Bradshaw, Samantha, and Philip N Howard. "Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation." Oxford, UK: Project on Computational Propaganda, December 2017.

Bradshaw, Samantha, and Philip N. Howard. "Challenging Truth and Trust: A Global Inventory of Organized Social Media Manipulation." Oxford, UK: Project on Computational Propaganda, 2018.

Bradshaw, Samantha, and Philip N. Howard. "The Global Disinformation Order: 2019 Global Inventory of Organised Social Media Manipulation." Oxford, UK: Project on Computational Propaganda, 2019.

Cairo, Alberto. "Ethical Infographics: In Data Visualization, Journalism Meets Engineering." IRE Journal 37 (2014).

Chandler, Daniel, and Rod Munday. A Dictionary of Media and Communication. 1st ed. Oxford Paperback Reference. Oxford ; New York: Oxford University Press, 2011.

Davies, Harry. "Ted Cruz Campaign Using Firm That Harvested Data on Millions of Unwitting Facebook Users." The Guardian, December 11, 2015, sec. US news.

Dubois, Elizabeth, and Fenwick McKelvey. "Canada: Building Bot Typologies." In Computational Propaganda: Political Parties, Politicians and Political Manipulation on Social Media. Oxford Studies in Digital Politics. New York, NY, USA: Oxford University Press, 2019.

Drucker, Johanna. "Humanities Approaches to Graphical Display." Digital Humanities Quarterly 005, no. 1 (March 10, 2011).

Foot, Kirsten A., and Steven M. Schneider. "Online Action in Campaign 2000: An Exploratory Analysis of the U.S. Political Web Sphere." Journal of Broadcasting & Electronic Media 46, no. 2 (June 2002): 222–44.

Foot, Kirsten A., Steven M. Schneider, Meghan Dougherty, Michael Xenos, and Elena Larsen. "Analyzing Linking Practices: Candidate Sites in the 2002

5. Bibliography

US Electoral Web Sphere." Journal of Computer-Mediated Communication 8, no. 4 (June 23, 2006): 0–0.

Glowacki, Monika, Vidya Narayanan, Sam Maynard, Gustavo Hirsch, Bence Kollanyi, Lisa-Maria

Neudert, Philip N Howard, and Thomas Lederer. "News and Political Information Consumption in Mexico: Mapping the 2018 Mexican Presidential Election on Twitter and Facebook." Oxford, UK: Project on Computational Propaganda, June 29, 2018.

Grassegger, Hannes. "Die Finkelstein Formel." Das Magazin, no. 1/2 (January 12, 2019): 9–17.

Griethe, Henning, and Heidrun Schumann. "Visualizing Uncertainty for Improved Decision Making,", 4th International Conference on Perspectives in Business Informatics Research (BIR'05), 2005.

Hedman, Freja, Fabian Sivnert, Bence Kollanyi, Vidya Narayanan, Lisa-Maria Neudert, and Philip N. Howard. "News and Political Information Consumption in Sweden: Mapping the 2018 Swedish General Election on Twitter." Oxford, UK: Project on Computational Propaganda, September 6, 2018.

Hepworth, Katherine. "Big Data Visualization: Promises & Pitfalls." Communication Design Quarterly Review 4, no. 4 (March 27, 2017): 7–19.

Hepworth, Katherine, and Christopher Church. "Racism in the Machine: Visualization Ethics in Digital Humanities Projects." Digital Humanities Quarterly 12, no. 4 (2018).

Hepworth, Katherine. "Ethical Visualization." https://github.com/kathep/ethics. (October, 2019).

Herring, Susan C. "Web Content Analysis: Expanding the Paradigm." In International Handbook of Internet Research, edited by Jeremy Hunsinger, Lisbeth Klastrup, and Matthew Allen, 233–49. Dordrecht: Springer Netherlands, 2010.

Howard, Philip N., Gillian Bolsover, Bence Kollanyi, Samantha Bradshaw, and Lisa-Maria Neudert. "Junk News and Bots during the U.S. Election: What Were Michigan Voters Sharing Over Twitter?" Oxford, UK: Project on Computational Propaganda, March 26, 2017.

Information Commissioner's Office. "Democracy Disrupted? Personal Information and Political Influence," July 11, 2018.

Information Commissioner's Office. "Investigation into Theuse of Data Analyticsin Political Campaigns: A Report to Parliament," November 6, 2018.

King, Gary, Jennifer Pan, and Margaret E. Roberts. "How the Chinese Government Fabricates Social Media Posts for Strategic Distraction, Not Engaged Argument." American Political Science Review 111, no. 3 (August 2017): 484–501.

Kosinski, Michal, David Stillwell, and Thore Graepel. "Private Traits and Attributes Are Predictable from Digital Records of Human Behavior." Proceedings of the National Academy of Sciences 110, no. 15 (April 9, 2013): 5802–5.

5. Bibliography

Krogerus, Mikael, and Hannes Grassegger. "Ich habe nur gezeigt, dass es die Bombe gibt." Tages-Anzeiger, March 20, 2018, sec. International.

L'Ecuyer, Pierre, Patrick Maillé, Nicolás E. Stier-Moses, and Bruno Tuffin. "Non-Neutrality of Search Engines and Its Impact on Innovation." Internet Technology Letters 1, no. 1 (January 2018): e10.

Lacy, Stephen, Brendan R. Watson, Daniel Riffe, and Jennette Lovejoy. "Issues and Best Practices in Content Analysis." Journalism & Mass Communication Quarterly 92, no. 4 (December 2015): 791–811.

Neudert, Lisa-Maria, Kollanyi Bence, and Philip N. Howard. "Junk News and Bots during the German Parliamentary Election: What Are German Voters Sharing over Twitter?" Oxford, UK: Project on Computational Propaganda, September 19, 2017.

Newton, Casey. "Facebook Gave Spotify and Netflix Access to Users' Private Messages." The Verge, December 18, 2018.

MacEachren, Alan M., Anthony Robinson, Susan Hopper, Steven Gardner, Robert Murray, Mark Gahegan, and Elisabeth Hetzler. "Visualizing Geospatial Information Uncertainty: What We Know and What We Need to Know." Cartography and Geographic Information Science 32, no. 3 (January 2005): 139–60.

Machado, Caio, Beatriz Kira, Gustavo Hirsch, Nahema Marchal, Bence Kollanyi, Philip N. Howard, Thomas Lederer, and Vlad Barash. "News and Political Information Consumption in Brazil: Mapping the First Round of the 2018 Brazilian Presidential Election on Twitter." Oxford, UK: Project on Computational Propaganda, October 5, 2018.

McCrae, Robert R., and Paul T. Costa. "Personality Trait Structure as a Human Universal." American Psychologist 52, no. 5 (1997): 509–16.

Püntener, Daniela. "Hotspot: Meine Datenspur" Hotspot, n.d. https://www.srf.ch/play/suche?query=hotspot%20datenspur

Pfeffer, Jürgen. "Visualization of Political Networks." In The Oxford Handbook of Political Networks. New York: Oxford University Press, 2017.

Popov, Vesselin. "How to Wield the Data-Driven Double-Edged Sword: Navigating the Ethics of Psychological Profiling with Big Data." ZKM | Zentrum für Kunst und Medien, Karlsruhe, Germany, April 12, 2017.

Reuter, Markus. "Social Bots: Was nicht erkannt werden kann, sollte nicht reguliert werden." netzpolitik.org (blog), May 9, 2019. https://netzpolitik.org/2019/social-bots-was-nicht-erkannt-werden-kann-sollte-nicht-reguliert-werden/.

Ruch, André. "«Wahlkampf – Digital»." 10vor10. SRF Schweizer Radio und Fernsehen, September 9, 2019. https://www.srf.ch/sendungen/10vor10/wahlkampf-digital.

Sanovich, Sergey, Denis Stukal, and Joshua A. Tucker. "Turning the Virtual Tables: Government Strategies for Addressing Online Opposition with an Application to Russia." Comparative Politics 50, no. 3 (April 1, 2018): 435–82.

5. Bibliography

Thomson, Judi, Elizabeth Hetzler, Alan MacEachren, Mark Gahegan, and Misha Pavel. "A Typology for Visualizing Uncertainty." edited by Robert F. Erbacher, Jonathan C. Roberts, Matti T. Grohn, and Katy Borner, 146. San Jose, CA, 2005.

Tufte, Edward R. Envisioning Information. 1st edition edition. Cheshire, Conn: Graphics Press, 1990.

Tufte, Edward R. The Visual Display of Quantitative Information. 2nd edition edition. Cheshire, Conn: Graphics Press, 2001.

Trice, Michael, and Liza Potts. "Building Dark Patterns into Platforms: How GamerGate Perturbed Twitter's User Experience." Present Tense - A Journal of Rhetoric in Society 6, no. 3 (January 21, 2018): 11.

Woolley, Samuel C., and Philip N. Howard. "Political Communication, Computational Propaganda, and Autonomous Agents - Introduction." International Journal of Communication 10 (2016) (2016): 4882–4890.

Zeifman, Igal. "2015 Bot Traffic Report: Humans Take Back the Web, Bad Bots Not Giving Any Ground | Imperva." Imperva Blog (blog), December 9, 2015. https://www.imperva.com/blog/bot-traffic-report-2015/.

Zhdanova, Mariia, and Dariya Orlova. "Ukraine: External Threats and Internal Challenges." In Computational Propaganda: Political Parties, Politicians and Political Manipulation on Social Media. Oxford Studies in Digital Politics. New York, NY, USA: Oxford University Press, 2019.