03 2018

Quantification Madness

An interview with Vladan Joler on the Facebook Algorithmic Factory report

Vladan Joler / Raimund Minichbauer

***Raimund Minichbauer**: In this Interview, we will talk about research conducted by Share Lab on Facebook.[1] Can you say a little bit about Share Lab and Share Foundation?*

Vladan Joler: We created Share Foundation in 2011. Initially it was a reaction to the positive ways of thinking about technology, the Internet, and network culture. We started organizing regular large-scale meetings of Internet activists and people related to net culture – gatherings, festivals, conferences – consisting of up to 2,000 participants. The aim was to give people the possibility of exchanging their views and ideas. However, we then realized that by just gathering people we were not able to substantially develop our capacity for dealing with problems. So we changed to a different strategy and created an organization, which was mainly based on expertise. We – in Belgrade and across Serbia – started to gather people with different kinds of expertise: lawyers, media theorists and people with a background in cyber forensics and different kinds of technical knowledge. This way we were able to transform into an organization capable of following various aspects

of information warfare – mainly attacks on investigative journalists, independent online media, etc. as they took place four or five years ago in Serbia and throughout the region. Regarding online attacks and placing pressure on journalists and media organizations, we were able to give them cyber-forensic support for trying to understand the attacks as well as assistance with legal expertise. We developed to an entity that could monitor all those cases, while, at the same time, acting as a kind of emergency "red cross" service in cyberspace. That was also a very interesting time in which Serbia was between two different regimes, and we could follow that in the cyber domain.

But when you run a big organization you need to devote a lot of time and energy to maintaining it – to developing a strategy for sustainability, applying for projects, etc. So some of us formed a group that wanted to be more active in the field of research. We started with investigations related to "invisible infrastructures," and we really started from the basics. We tried to understand what different networks look like. We then tried to map different kinds of data flows. Through this we delved deeper and deeper into analyzing something we like to refer to as "surveillance capitalism." We started to research different kinds of trackers, different kinds of surveillance technologies, and different companies that were active in the field. That led us to the investigation of one organization, called "Hacking Team."[2] We tried to use similar methodologies to the ones that NSA probably uses. We collected metadata from e-mails, etc., but for tracking people and organizations from the other side – the "bad guys" so to speak. We felt like we were some kind of detectives or investigative journalists, but a strange kind, because we investigated machines, tools and processes, we followed data, and so on. That led us to this investigation on Facebook and to the various algorithms that shape our reality through the

moderation of information feeds, and to the analysis of the processes of how we are transformed from users to products, and how they basically transform our online behavior into power and profit.

RM: What were the basic questions or general hypotheses from which you started the Facebook research?

VL: We did not have a specific goal with the research. Most of our investigations are just led by the idea that we are walking in the dark and we do not know a lot about what is happening behind screens and behind the technology that we use. Our main motivation has been to switch the light on. It was the same with the research on Facebook: Let us try to find out how we, from the position of being outside of the box, can be active in the sense of algorithmic transparency. We wanted to test our capacity for conducting a kind of independent audit of such a complex system.

RM: Before we go into the details of your research, could you briefly please explain the essential concept behind the "social graph"?

VJ: The social graph is the heart of the system and of the ontology that they are building. It is the map of all of the actors and objects and relations within the system. There are hundreds of different ways in which the algorithms extract data. Each of the pieces of data that have been extracted become part of this one huge map. Each time we upload something or perform some other activity, it becomes one node in this system. This node can be related to another node, for instance, when I upload a picture, that picture becomes a node. However, I, the user, am another node. And the relation between myself and the picture in this example is "uploaded" – user / uploaded / node. Other actors in the graph can

have different relations with the same node – like, share, tag or whatever they can do with that one picture. The social graph is a multidimensional map/graph of everything that is inside this "empire," and they are able to perform different kinds of algorithmic analyses and statistical analyses on top of that map.

RM: Is there a fundamental difference between a user-object and e.g. a photo-object within this map or ontology of the social graph or is it basically the same thing?

VJ: On this level, they are the same. Some objects can have a larger number of different forms of relations with other objects, but everything is an object. There is no hierarchy of objects. The graph is the core of the process, and then different algorithms try to draw different forms of conclusions from the map. This can go in different directions, like serving you a targeted ad or understanding to which social class you belong or presenting you the news feed. There are hundreds of different algorithms, extracting different kinds of conclusions and knowledge from this map. In our research, we followed one stream – how different algorithms form the product at the end of this process, the target user profiles.

RM: When we follow your analyses from collecting via processing to targeting users, one gets the impression that the individual user is sort of dissolved in a fog of data and then reconstructed as an individual entity at the end of the process. Would you agree with this interpretation?

VJ: I do not have the feeling that the reconstruction of us as individuals at the end of the process is so important. Sometimes we are treated as a group. Sometimes we are treated as individual targets. The processes of understanding us and what our behaviors mean and how these behaviors define us is matched in some kind

of fuzzy logic with the processes of understanding the other side, the advertisements and what the customer wants, who is buying our profile, and buying us as a target. This does not relate to the level of the individual, but to the level of groups of people that are defined by various characteristics. One such group may be formed of people with rather differing characteristics: People who do not just like the same kind of music, but different kinds; people who do not just live in the same part of the city, but in different parts. However, for the system, this is *one* group. Let's call this group *107888179*. The logic of how the machine comes up with the idea that this is *one* group is completely artificial. It is probably not possible for us human spectators to understand why those people belong to the same group. When we look at the processes in creating target groups, only the first layer is directly defined by human logic, for instance: targeting based on gender. This is something that we can understand, but how groups are formed on a deeper level is rather artificial. An example to illustrate this might be Amazon's automated warehouses where the articles that have been ordered are packed and dispatched. There are lots of shelves, and these shelves are attached to robots. A product is put in a specific place on a shelf, but the place in which they are put is not according to some human logic like it would be with various watches being put on a shelf. No, it is defined by algorithms, by live feeds from what people are buying, what they buy together in one order, etc. A human worker who works in this space would not be able to find anything without being guided by the software that is also connected to this algorithmic system. I think it is the same thing on some level of Facebook and their algorithms – it goes beyond human logic.

RM: As you explained, there are aggregations on collective levels in the processing of data. Segments and groups are constructed. I

would have expected that the active interactions between people – friends, groups that form – would be of more importance in a social network. But when I read your reports I had the impression that it is still much more about who likes red wine and Harry Potter, to use the example from your writing. And my connection to another human user is mainly interpreted in answering the question of whether the other user also likes red wine, which would support the hypothesis that I belong to that target group. It seems that these are just "taxonomic collectives," meaning that they are purely constructed through observation, and the interactions between people are far less important than I expected.

VJ: Yes, I think we can put it that way. When we approach it from the level of human understanding: For us as human users, it is usually most important that we belong to a group of friends and that we are part of this community and we have a way of understanding our social relations. But it is different for algorithms. For example, what you see in the news feed is defined by some main algorithm as well as by hundreds of thousands of different data points. The algorithms are able to penetrate some patterns of anomalies in our behavior which are only important to "them" when they try to understand what we do. But for us as human beings this would be complete nonsense. Moreover, when they try to understand the social classes that we belong to, for example, there are different kinds of calculations that they perform, and those go beyond the usual categories of financial status, the part of the city in which we live, etc. For them, there are many different data points that can be used to define the social classes to which we belong.

RM: As quantification is Facebook's main form of production, what is the impact of this on user behavior or on the behavior that

the interface enables and supports? Reducing qualitative differences is the first step in quantification ...

VJ: I think that this is basically "quantification madness." You can follow this back to the very beginnings of the scientific method. We as human beings try to quantify everything – living and unliving nature, our behavior, our emotions, and so on. I mean this in the sense of Lyotard's "affinity to infinity." Contemporary capitalism has the means to explore, conquer, and colonize different frontiers, and with the technologies that we have today these frontiers can go on indefinitely. Now we are into the quantification of our affective and cognitive behavior. This is a new field of exploitation, and they are trying to extract as much as they can from their resources – which are basically *us*. This is what Facebook does with quite a bit of success. Then they use that for profit creation, the accumulation of power, etc.

RM: When Facebook developers design some part of the interface they are aware that the aim behind all of it is quantification. When you analyzed Facebook's patents, did you have the impression that this aim for quantification was creating some kind of structuring or mechanization of user behavior? For example, I recently read an article on social bots, which stated that it was not all too difficult for social bots to mimic the behavior of human users in social networks, because human activities are already very mechanized.

VJ: When we talk about interfaces, I think that the differences between Facebook and its "predecessor" MySpace provide a good example. In MySpace, you had the freedom to construct your interface to a certain degree, to create different things, to design how your page looks, etc., but in Facebook everything is structured. For us as immaterial workers, it is structured like cubicles. There is not a lot that you can do or create other than

that which is meant for you to do: post something, comment, etc. There is no way of going beyond the given field. Those interfaces have become a good method of control. They control how people behave and in which ways they produce things on those platforms. In regard to that control, during our research process, we also became aware of the fact that platforms like Facebook do not merely exist for the expression of our emotions, social relations, etc., but that they are increasingly used, for example, by the department of homeland security, which checks your Facebook page when you enter the United States, or by insurance agencies, banks, etc. When we as users understand that, we are pushed to behave differently. From the moment that we know that someone is watching us, and that what we say and which pictures we upload can influence our credit scores or whether we will be able to enter the United States, we need to change our behavior. Through these methods, these platforms become places of control. We stop freely expressing our emotions and we start maintaining pages and profiles as something that should create a nice public appearance. We modify our behavior in different ways. I think that this is a threat to Facebook itself, because their main resource is our emotions and our behavior, and if these are repressed, Facebook does not get enough quality resources. Then they start getting rather clean professional profiles instead of profiles filled with emotions, likes or whatever else.

RM: Do you think that these concerns are also behind the announcements that they made at the beginning of 2018, which said that Facebook would go back to its roots with people-to-people connections being placed at the center, thereby becoming a kind of nice, cozy social network?

VJ: It is, because they are worried. The dissatisfaction with Facebook has been growing over the last half a year to a year. And

now the critiques that Google and Facebook are evil, that they destroy the social tissue, etc. have been becoming mainstream narratives. I think this is their concern now – that they may stop being a cozy place and become a place for public profiles, like LinkedIn. That does not mean that they will have a problem with the exploitation of data in general, but that the exploitation will become shallower than it is now.

RM: Facebook is not only continually extracting data from its user base, it also conducts experiments. This was not part of your research, but you probably checked out what you could have found out or learned from it, right?

VJ: In one experiment conducted very recently, Serbia was one of six countries where Facebook tested changes to its News Feed.[3] These are now implemented in the "less media more friends" policy. This, of course, is not the first time that they have performed experiments. I think the entire process of managing the network is based on smaller experiments, on trying to tweak the system a bit to try to understand what generates more or less profit and how people change their behaviors in reaction to certain modifications. When talking about terrorist attacks, there is of course a difference between how you filter such information, whether you give people more pictures of happy dogs and cats or more pictures of blood and dead people. And I am sure that Facebook performs a lot of experiments like these in different groups and segments, and on national levels, etc.

RM: To which extent is it possible to obtain Facebook data and to perform research from outside of the network?

VJ: Our capacity for investigation is very limited. There are different reasons for that: First of all, the system is very complex.

We initially planned to do some kinds of measurements. Then we realized that it was not possible to have a clean environment for experimentation. From the very moment you go and visit a website, you have already contaminated the experiment. And there are so many different data inputs that you really cannot understand or know which one influences what. Furthermore, it is very hard to investigate those black boxes on the level of data or to try and understand that data or to reverse engineer it. There are experiments of various scales and tools that have been developed from the outside, but they are only able to understand a small portion of the huge mosaic. For such an investigation you would also need a lot of resources and human capacity – for instance, data analysts and cyber forensic scientists, data scientists, AI scientists, etc. We will thus never be able to compete with a platform like Facebook to this degree. Because of their financial power, they are able to buy the best minds that come out of the universities, and no independent group or even government can compete with them. I am also skeptical about the very idea of algorithmic transparency. Even if we were able to understand the process from the outside and to reverse engineer it in some way, it would not work, because the system has constantly been changing in the time that you spent doing that, and it will have developed more algorithms and will have experimented with new things in the meantime. For example, the map that we produced is more of a symbolic map rather than a precise, accurate one, because many different parts of the map belong to different patterns that existed at different times. In reality, there has probably never been a moment in which the process really looked like this. Still, that's the map we have and even if it is not precise or accurate, it is something that you can look at on more of a symbolic level for trying to understand complexity rather than on the mere level of facts. Another problematic point concerning the idea of

algorithmic transparency is the question: *For whom should it be transparent?* If you do an online search for "Facebook algorithms," the first three to five pages of search results will mainly be entries from marketing agencies that try to hack Facebook algorithms in order to do their own thing. If the Facebook algorithms were fully transparent, it would be easier for companies like Cambridge Analytica,[4] for example, to misuse this process or for it to be misused in order to target people during elections, and so on. It is a tricky question – how processes can be transparent and in which way.

RM: Departing from the research you have done, what do you think about the various forms of possible resistance?

VJ: In the sphere of social media critique, we can say that there are three main approaches: The first one attempts to go beyond the field of technology. This can refer to Neo-Luddist ideas of rejecting technology, of going out into nature and trying to find a new balance between ourselves, nature, and technology. The second one is staying in – e.g. by trying to regulate the space. This includes different kinds of initiatives regarding how governments, such as the EU, regulate Facebook. It can also refer to staying in but trying to develop different relations between *us* as users and *them* as the owners of those factories. For instance, this can include trying to organize unions of immaterial workers or different kinds of ideas that open a dialog between user-workers and owners. There are also more radical ideas like obfuscation, but these ultimately remain within the system. For example, one can have a Facebook account, but they can find ways of confusing the system. These are the people who want to remain and fight within the factory. The third approach is to fundamentally change or rebuild the system. The idea is that e.g. social networks are not wrong in and of themselves, but we need new and different social networks.

So if we rebuild them with open source tools, etc., things will be alright. I am not so sure which one is the best.
There are two interesting cases from the past: The first one is GeoCities, which had a near monopoly on hosting personal web pages in the late 1990s / early 2000s. The second one is Friendster, which had a position similar to that of Facebook but before Facebook developed. Both of them had similar roles to the one Facebook has today. There is a kind of forensic analysis of the death of Friendster.[5] Researchers have been following its death on the level of the social graph in order to analyze how Friendster's social graph collapsed. It started from the periphery, from the small "unimportant" nodes that started to leave. Through this process, the main nodes became more isolated from one another, because the small nodes were the links between the main nodes. This process is similar to the expansion and cooling of the universe in which all of the galaxies become more and more isolated from one another. This is what happened to Friendster. The main nodes lost their social capital and the system collapsed. I think this is a possible scenario for the death of Facebook – collapsing from within, not from the outside. Another scenario is that it would stop producing enough profit for its owners. At that moment, the plug would be pulled. Then we come to the question of resources. We are the resources and our emotions are the resources. If we want to stop the system from working, we should stop supplying it with resources. We should stop producing emotions there.

RM: How did GeoCities end?

VJ: It was related to the financial situation. Yahoo bought GeoCities for I don't know how many billions of dollars. Then there was the dot-com crash. Then Yahoo was not able to define their business strategy for a while, and GeoCities collapsed because they did not fulfill their business promises.

RM: Share Lab has dealt with privacy issues before, e.g. there is a piece on what we can learn about a person from their browser history. It seems that the concept of privacy on an individual level has its limitations ...

VJ: A few years ago, there was a certain period of time when the question of privacy was the core question for many organizations, including our own. But the more you dive into these issues, you realize that the possibilities for exploiting our behavior in relation to technology are limitless. They can go really deep. Even a simple thing like electricity flows can teach us a lot about people plugging in and unplugging their devices. And this is just surveillance on the level of electrons going through wires. It can go deeper and deeper. We will never succeed in defending privacy.
Also, our definitions of privacy are constantly changing. It is not the same now as it was fifty or one hundred years ago, and it will not be the same next year. This is an ongoing process, which is never going to achieve full quality, but it is important to continue this discussion and to try and understand what privacy means in the context of technology and networks as well as in the current discussions on artificial intelligence. I think it is interesting as a field of ethical discussion, as something that we are trying to achieve, as a quality, but it is really hard to get there. Furthermore, the discussion on privacy reached its peak a few years ago, and has been in a decline since then. This is also part of the more general problem of us switching from one hot topic to the next really quickly. Four years ago, it was privacy. Then it was algorithmic transparency. Now it is artificial intelligence. We are constantly switching from one hot topic to another hot topic without solving much along the way.

RM: You already mentioned that the critique of Facebook reached the mainstream following the problems surrounding the US

elections, fake news, as well as concerns regarding generations of increasingly addicted and distracted children. Do you think that this can help to create a new situation?

VJ: It already is a new situation. The critique, which has been around in specialized discourses for a while, e.g. mailing lists like *nettime* or *unlike-us*, has now hit the mainstream. For example, a few days ago, you had a talk by George Soros[6] on how Facebook and Google are a menace to society. I am really curious how this will influence positions on Facebook and Google. It seems that we have entered a time of rethinking their positions. We are far from having techno-utopian faith in them now.

January 2018
Language editing: Lina Dokuzovic
The interview is part of *Midstream* (http://midstream.eipcp.net/interviews).

[1] This Interview focuses on the *Facebook Algorithmic Factory* report, which was published online in three parts: "Immaterial Labour and Data Harvesting" (https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/), "Human Data Banks and Algorithmic Labour" (https://labs.rs/en/facebook-algorithmic-factory-human-data-banks-and-algorithmic-labour/), and "Quantified Lives on Discount" (https://labs.rs/en/quantified-lives/). Two additional texts on Facebook can be found at: https://labs.rs/en/category/facebook-research/.

[2] https://en.wikipedia.org/wiki/Hacking_Team

[3] See: http://www.bbc.com/news/technology-41733119, https://www.nytimes.com/2017/11/15/opinion/serbia-facebook-explore-feed.html

[4] https://en.wikipedia.org/wiki/Cambridge_Analytica

[5] David Garcia, Pavlin Mavrodiev, Frank Schweitzer, "Social Resilience in Online Communities: The Autopsy of Friendster," 2013: https://arxiv.org/pdf/1302.6109.pdf

[6] https://www.theguardian.com/business/2018/jan/25/george-soros-facebook-and-google-are-a-menace-to-society