35. Algorithms in the Spotlight: Collaborative Investigations at *Der Spiegel*

Christina Elmer

Abstract

Using several examples from German media, this chapter provides insights into collaborative projects that investigate algorithms and summarizes the most important lessons learned.

Keywords: algorithms, algorithmic accountability, data journalism, crowdsourcing, transparency, collaboration

The demand for transparency around algorithms is not new in Germany. In 2012, *Der Spiegel* columnist Sascha Lobo called for the mechanics of the Google search algorithm to be disclosed (Lobo, 2012), even if this would harm the company. The reason was that Google can shape how we view the world, for example through the autocomplete function, as a prominent case in Germany illustrated. In this case, the wife of the former federal president had taken legal action against Google because problematic terms were suggested in the autocomplete function when her name was searched for. Two years later, the German minister of justice repeated this appeal, which was extended again by the federal chancellor in 2016: Algorithms should be more transparent, Angela Merkel demanded (Kartell, 2014; Reinbold, 2016).

In the past few years, the topic of algorithmic accountability has been under constant discussion at *Der Spiegel*—but initially only as an occasion for reporting, not in the form of our own research or analysis project. There may be two primary reasons why the German media began experimenting in this area later than their colleagues in the United States. First, journalists in Germany do not have such strong freedom of information rights and

instruments at their disposal. Second, data journalism does not have such a long tradition as in the United States. *Der Spiegel* has only had its own data journalism department since 2016 and is slowly but steadily expanding this area. It is, of course, also possible for newsrooms with smaller resources to be active in this field—for example, through cooperation with organizations or freelancers. In our case, too, all previous projects in the area of algorithmic accountability reporting have come about in this way. This chapter will therefore focus on collaborations and the lessons we have learned from them.

Google, Facebook and Schufa: Three Projects at a Glance

Our editorial team primarily relies on cooperation when it comes to the investigation of algorithms. In the run-up to the 2017 federal elections, we joined forces with the NGO AlgorithmWatch to gain insights into the personalization of Google search results.¹ Users were asked to install a plug-in that regularly performed predefined searches on their computer. A total of around 4,400 participants donated almost six million search results and thus provided the data for an analysis that would challenge the filter bubble thesis—at least regarding Google and the investigated area.

For this project, our collaborators from AlgorithmWatch approached *Der Spiegel*, as they were looking for a media partner with a large reach for crowdsourcing the required data. While the content of the reporting was entirely the responsibility of our department covering Internet- and technology-related topics, the data journalism department supported the planning and methodological evaluation of the operation. Furthermore, the backup of our legal department was essential in order to implement the project in a way which was legally bulletproof. For example, data protection issues had to be clarified within the reporting and had to be fully comprehensible for all participants involved in the project.

Almost at the same time, *Der Spiegel* collaborated with *ProPublica* to deploy their AdCollector in Germany in the months before the elections (Angwin & Larson, 2017). The project aimed to make transparent how German parties target Facebook users with ads. Therefore, a plug-in collected the political ads that a user sees in her stream and revealed those ads that are not displayed to her. For this project, *Der Spiegel* joined forces with other German media outlets such as *Süddeutsche Zeitung* and *Tagesschau*—an

1 <https://algorithmwatch.org/filterblase-geplatzt-kaum-raum-fuer-personalisierung-bei-google-suchen-zur-bundestagswahl-2017/> (German language)

unusual constellation of actors who usually are in competition with each other. In this case it was necessary to reach as many people as possible to serve the public interest. The results could also be published as journalistic stories, but our primary focus was transparency. After two weeks, around 600 political advertisements had been collected and made available to the public.

ProPublica's Julia Angwin and Jeff Larson introduced the idea of a collaboration at the annual conference of the German association of investigative journalists, Netzwerk Recherche in Hamburg, where they held a session on algorithmic accountability reporting. The idea was developed from the very beginning in collaboration with technical and methodology experts from multiple departments in the newsroom of *Der Spiegel*. The exchange with our previous partner, the non-profit AlgorithmWatch, was also very valuable for us in order to shed light on the legal background and to include it in our research. After the conference, we expanded the idea further through regular telephone conferences. Our partners from the other German media outlets became involved at later stages as well.

In 2018, *Der Spiegel* contributed to a major project to investigate an extremely powerful algorithm in Germany—the Schufa credit report. The report is used to assess the creditworthiness of private individuals. The report should indicate the probability that someone can pay their bills, pay the rent or service a loan. It can therefore have far-reaching implications for a person's private life and a negative effect on society as a whole. For example, it is conceivable that the score may increase social discrimination and unequal treatment of individuals, depending on the amount of data that is available about them. Incorrect data or mix-ups could be fatal for individuals. The algorithm's underlying scoring is not transparent. Which data is taken into account in which weighting is not known. And those affected often have no knowledge of the process. This makes Schufa a controversial institution in Germany—and projects like OpenSCHUFA absolutely vital for public debate on algorithmic accountability, in our opinion.²

The project was mainly driven by the NGOs Open Knowledge Foundation (OKFN) and AlgorithmWatch. *Der Spiegel* was one of two associated partners, together with Bayerischer Rundfunk (Bavarian Broadcasting). The idea for this project came up more or less simultaneously, with several parties involved. After some successful projects with the NGOs AlgorithmWatch and

2 <http://www.openschufa.de/> (German language)

OKFN as well as with the data journalism team of Bayerischer Rundfunk, *Der Spiegel* was included in the initial discussions.

The constellation posed special challenges. For the two media teams, it was important to work separately from the NGOs in order to ensure their independence from the crowdfunding process in particular. Therefore, although there were, of course, discussions between the actors involved, neither an official partnership nor a joint data evaluation were possible. This example emphasizes how important it is for journalists to reflect on their autonomy, especially in such high-publicity topics.

Making OpenSCHUFA known was one of the central success factors of this project. The first step was to use crowdfunding to create the necessary infrastructure to collect the data, which was obtained via crowdsourcing. The results were jointly evaluated by the partners in the course of the year in anonymized form. The central question behind it is: Does the Schufa algorithm discriminate against certain population groups, and does it increase inequality in society? According to the results, it does. We found that the score privileged older and female individuals, as well as those who change their residence less frequently. And we discovered that different versions of algorithms within the score generated different outcomes for people with the same attributes, a type of discrimination that was not previously known regarding this score.

These results would not have been possible without the participation of many volunteers and supporters. The crowdfunding campaign was largely successful, so that the financing of the software could be secured within the framework.³ And within the subsequent crowdsourcing process, about 2,800 people sent in their personal credit reports. This sample was, of course, not representative, but nevertheless sufficiently diverse to reveal the findings described.

Impact and Success Indicators

Both the Facebook and the Google investigations were rather unspectacular in terms of findings and confirmed our hypotheses. Political parties apparently hardly used Facebook's targeting options and the much-cited Google filter bubble was not found in our crowdsourcing experiment in Germany. But for us the value of these projects lay in increasing our readers' literacy around functionalities and risks of algorithms in society.

3 <https://www.startnext.com/openschufa> (German language)

The reach of our articles was an indicator that we had succeeded in making the topic more widely known. The introductory article at the start of the Schufa project reached a large audience (around 335,000 readers).⁴ The reading time was also clearly above the typical one—with an average of almost three minutes. In addition, the topic was widely discussed in public arenas and covered by many media outlets and conferences.

The topic has also been debated in political circles. After the publication of the Schufa investigations, the German minister of consumer protection called for more transparency in the field of credit scoring. Every citizen must have the right to know which essential features have been included in the calculation of their creditworthiness and how these are weighted, she demanded.

What about impact on everyday reality? As a first step, it was important for us to contribute to establishing the topic in the public consciousness. So far, we have not seen any fundamentally different way political actors deal with algorithms that have broader societal consequences.

Nevertheless, the topic of algorithmic accountability reporting is very important to us. This is because in Europe we still have the opportunity to debate the issue of algorithms in society and to shape how we want to deal with it. It is part of our function as journalists to provide the necessary knowledge so that citizens can understand and shape the future of algorithms in society. As far as possible, we also take on the role of a watchdog by trying to make algorithms and their effects transparent, to identify risks and to confront those responsible. To achieve this, we have to establish what might otherwise be considered unusual collaborations with competitors and actors from other sectors. We hope that such alliances will ultimately increase the pressure on legislation and transparency standards in this area.

More effort and resources need to be dedicated to algorithmic accountability investigations and “The Markup” has published some very exciting research in this area. Further experimentation is very much needed, partly because there is still scope for action in the regulation of algorithms. The field of algorithmic accountability reporting has only begun to develop in recent years. And it will have to grow rapidly to meet the challenges of an increasingly digitized world.

4 The majority of these, however, came to the article via internal channels like our homepage. This was different in the case of another article, the field report featuring the author's personal story, which was read by around 220,000 people. A fifth of them reached the article via social media channels, which is well above the average. So it seems that we were able to reach new target groups with this topic.

Organizing Collaborative Investigations

Running collaborative investigations takes a whole set of new or less used skills in the newsroom. This includes the analysis of large data sets and the programming of specific procedures, but also the management of projects. The latter is too easily overlooked and will be described in more detail here, with concrete examples from our previous work.

Working together in diverse constellations not only makes it easier to share competencies and resources, it also allows a clear definition of roles. As a media partner, *Der Spiegel* positioned itself in these collaborations more as a neutral commentator, not too deeply involved in the project itself. This allowed the editors to remain independent and thus justify the trust of their readers. They continued to apply their quality criteria to reporting within the project—for example, by always giving any subject of their reporting the opportunity to comment on accusations. Compared to the NGOs involved, these mechanisms may slow media partners down more than they are comfortable with, but at the same time they ensure that readers are fully informed by their reports—and that these will enrich public debate in the long term.

Reaching agreement about these roles in advance has proven to be an important success criterion for collaborations in the field of algorithmic accountability. A common timeline should also be developed at an early stage and language rules for the presentation of the project on different channels should be defined. Because, after all, a clear division of roles can only work if it is communicated consistently. This includes, for example, a clear terminology on the roles of the different partners in the project and the coordination of disclaimers in the event of conflicts of interest.

Behind the scenes, project management methods should be used prudently, project goals should be set clearly and available resources have to be discussed. Coordinators should help with the overall communication and thus give the participating editors the space they need for their investigations. To keep everyone up to date, information channels should be kept as simple as possible, especially around the launch of major project stages.

Regarding editorial planning, the three subject areas were challenging. Although in general relevance and news value were never questioned, special stories were needed to reach a broad readership. Often, these stories focused on the personal effects of the algorithms examined. For example, incorrectly assigned Schufa data made it difficult for a colleague from the *Der Spiegel* editorial team to obtain a contract with an Internet provider. His experience report impressively showed what effects the Schufa algorithm

can have on a personal level and thus connected with the reality of our audience's lives (Seibt, 2018).

Thus, we tailored the scope of our reporting to the interests of our audience as far as possible. Of course, the data journalists involved were also very interested in the functioning of the algorithms under investigation—an interest that is extremely useful for research purposes. However, only if these details have a relevant influence on the results of the algorithms can they become the subject of reporting—and only if they are narrated in a way that is accessible for our readers.

Internally in the editorial office, support for all three projects was very high. Nevertheless, it was not easy to free up resources for day-to-day reporting in the daily routine of a news-driven editorial team—especially when the results of our investigations were not always spectacular.

Lessons Learned

By way of conclusion, I summarize what we have learned from these projects.

Collaborate where possible. Good algorithmic accountability investigations are only possible by joining forces with others and creating teams with diverse skill sets. This is also important given both the scarcity of resources and legal restrictions that most journalists have to cope with. But since these projects bring together actors from different fields, it is crucial to discuss beforehand the underlying relevance criteria, requirements and capabilities.

Define your goals systematically. Raising awareness of the operating principles of algorithms can be a first strong goal in such projects. Of course, projects should also try to achieve as much transparency as possible. At best we would check whether algorithms have a discriminatory effect—but project partners should bear in mind that this is a more challenging goal to attain, one that requires extensive data sets and resources.

Exercise caution in project implementation. Depending on the workload and the day-to-day pressure of the journalists involved, you might even need a project manager. Be aware that the project timeline may sometimes conflict with reporting requirements. Take this into account in communicating with other partners and, if possible, prepare alternatives for such cases.

Invest in research design. To set up a meaningful design that produces useful data, you might need specialized partners. Close alliances with scientists from computer science, mathematics and related disciplines are particularly helpful for investigating some of the more technical aspects of algorithms. Furthermore, it may also be useful to cooperate with social

and cultural researchers to gain a deeper understanding of classifications and norms that are implemented in them.

Protect user data. Data donations from users may be useful to investigate algorithms. In such crowdsourcing projects legal support is indispensable in order to ensure data protection and to take into account the requirements of the national laws and regulations. If your company has a data protection officer, involve them in the project early on.

Works Cited

- Angwin, J., & Larson, J. (2017, September 7). Help us monitor political ads online. *ProPublica*. <https://www.propublica.org/article/help-us-monitor-political-ads-online>
- Kartell, S. vor. (2014, September 16). Maas hätte gerne, dass Google geheime Suchformel offenlegt. *Der Spiegel*. <https://www.spiegel.de/wirtschaft/unternehmen/google-heiko-maas-fordert-offenlegung-von-algorithmus-a-991799.html>
- Lobo, S. (2012, September 11). Was Bettina Wulff mit Mettigeln verbindet. *Der Spiegel*. <https://www.spiegel.de/netzwelt/netzpolitik/google-suchvorschlaege-was-bettina-wulff-mit-mettigeln-verbindet-a-855097.html>
- Reinbold, F. (2016, October 26). Warum Merkel an die Algorithmen will. *Der Spiegel*. <https://www.spiegel.de/netzwelt/netzpolitik/angela-merkel-warum-die-kanzlerin-an-die-algorithmen-von-facebook-will-a-1118365.html>
- Seibt, P. (2018, March 9). Wie ich bei der Schufa zum “deutlich erhöhten Risiko” wurde. *Der Spiegel*. <https://www.spiegel.de/wirtschaft/service/schufa-wie-ich-zum-deutlich-erhoehten-risiko-wurde-a-1193506.html>

About the Author

Christina Elmer established the data journalism section and fosters innovation projects in the editorial R&D team at *Der Spiegel* and is a board member of Netzwerk Recherche.