

# ALGORITHMS, FILTER BUBBLES AND FAKE NEWS

BY DOMINIK GRAU, Managing Director, Beuth

Edited by Jessica Patterson

## Hotsheets

Below the surface of digital platforms lies a deep sea of algorithms. Hidden in plain sight, they shape how we consume content and the way content or services are recommended to us. Publishers need to understand the growing influence of algorithms and the digital filter bubbles they create.

### THE NATURE OF ALGORITHMS

Algorithms help publishers, platforms and brands to amplify and personalize content on almost all digital channels. Most of today's content management systems and recommendation engines are driven by tens of thousands of algorithms. On the most basic level, algorithms are "sets of defined steps structured to process instructions/data to produce an output" according to National University of Ireland Professor Rob Kitchin. Kitchin, who recently published a paper on 'Thinking critically about and researching algorithms,' states that "algorithms can be conceived in a number of ways – technically, computationally, mathematically, politically, culturally, economically, contextually, materially, philosophically, ethically."<sup>1</sup> Dartmouth professors Tom Cormen and Devin Balkcom recently offered a specific example of the binary nature of algorithms: "Binary search is an efficient algorithm for finding an item from a sorted list of items. It works by repeatedly dividing in half the portion of the list that could contain the item, until you've narrowed down the possible locations to just one."<sup>2</sup>

Just like mathematical formulae, algorithms compute predefined metrics and calculations, step by step. Their code has no bias, they "know" or "understand" nothing, yet they compute everything that has been implemented as a rule or path. Every algorithm exists because there was an initial human operation during its creation. In the best-case scenario, for publishers and readers alike, algorithms behave like good human editors. University of Amsterdam professor Judith Möller writes, "The task of algorithms in the context of news recommendation systems is to structure and order the pool of available news according to predetermined principles." Three of the most common recommendation approaches according to Möller are:

1. **General popularity**, of the item, where all users get the same recommendation; popular items become even more popular and unpopular items disappear.
2. **Semantic filtering**, where items match the currently used one or previously used items.
3. **Collaborative filtering** or social information filtering, where 'word-of-mouth' recommendations are automatized over time.

Möller concludes that "the random recommender produces one of the most diverse recommendation sets" but that the "recommendations produced by the editors are not particularly diverse. If data on past user preferences are not taken into account, collaborative filtering produces the least amount of diversity".

### FROM BAD ALGORITHM TO FAKE NEWS

How does what we have learned about algorithms affect our daily online experience beyond reading pre-filtered news and stories? Let us assume that you have seen a product recommendation for a new smartphone on a social network, posted by a friend. He shared a link from an e-store. When you click on that link, the social platform will – if you use the default security settings – store cookies on your computer with dozens of data items directly assigned to you, such as time, location, computer

---

<sup>1</sup> <http://futuredata.stanford.edu/classes/cs345s/handouts/kitchin.pdf>

<sup>2</sup> <https://www.khanacademy.org/computing/computer-science/algorithms/intro-to-algorithms/v/what-are-algorithms>

settings, demographic assumptions, purchase behaviors and platform identifiers. You then look at the smartphone page on the e-store, which will store additional cookies on your computer.

Once you return to the social network, you might see multiple products ads for smartphones and accessories – either from the e-store or third-party websites who use redirection tools on most social platforms to track potential customers. If you continue your journey by comparing smartphone prices on search engines, even more cookies will be stored.

In some cases, the ad networks serve us ads for months on participating websites that use the ad platform. While we see personalized ads because of cookies and search history, we might also see content recommendations that are tailored to our interests. All the data is derived from demographic assumptions based on cookies or similar technologies. This is where the most basic form of a filter bubble of our own making begins to take shape. In case you visit dozens of websites every week – like most users do nowadays – dozens of potential filter bubbles emerge.

Another example of a situation where algorithms for personalized content, auto-recommendations and accompanying online ads had more severe consequences was in 2016, according to Dante Barone from Brazilian Universidade Rio Grande do Sul,<sup>3</sup> when the private company Cambridge Analytica combined neutral data mining algorithms and data analysis with biased calculations to customize and publish fake news related to Brexit and the U.S. presidential election.

The general public, who mostly does not know how or why media websites use algorithms, only learned about this manipulative strategy after Brexit and the U.S. election happened. This comes as no surprise as research shows that most of the U.S.-based students surveyed did not know how and why news websites use editorial judgments and data mining to deliver personalized results, according to Elia Powers from Towson University.<sup>4</sup> These two factors, the hidden use of algorithms by Cambridge Analytica to deliver manipulated content and that readers do not have the skills to understand how such fake news engines work, present an increasingly dangerous situation for publishers.

Increasing disinformation campaigns and false information makes it more difficult for “real” news to cut through the noise. A representative and wide-ranging German study by Christian Reuter, Katrin Hartwig, Jan Kirchner and Noah Schlegel, published in 2019 for the Technical University of Darmstadt, showed that 80 per cent of participants agree that fake news poses a threat and half of the respondents noticed fake news recently.<sup>5</sup>

There hardly is a way for readers to escape this phenomenon – unless publishers educate them on how to escape the influence of algorithms and filter bubbles.

## HOW PUBLISHERS CAN COUNTER FAKE NEWS

“As web companies strive to tailor their services (including news and search results) to our personal tastes, there's a dangerous unintended consequence: We get trapped in a filter bubble and don't get exposed to information that could challenge or broaden our worldview”, said Upworthy chief executive Eli Pariser in a TED talk on filter bubbles and information exposure.<sup>6</sup> Countering fake news therefore requires a deep understanding of algorithms and filter bubbles. The first steps for publishers in dealing with this challenge is an efficiently structured approach that considers how fake news is created and circulated.

In her research on fake news, Bente Kalsnes, an associate professor of communication from Oslo Metropolitan University, differentiates between efforts to counter fake news “directed toward legal, financial, and technical aspects to individuals’ media and information literacy and new fact-checking services.” She concludes that while disinformation is not a new phenomenon, “new communication technologies have made it easier than ever to produce and distribute falsehood and lies, dressed up as news to gain trustworthiness. The societal concerns that disinformation raises are numerous. New tools and methods—including

<sup>3</sup> [https://www.incobra.eu/object/news/714/attach/Fake\\_News\\_and\\_\\_Artificial\\_Intelligence.pdf](https://www.incobra.eu/object/news/714/attach/Fake_News_and__Artificial_Intelligence.pdf)

<sup>4</sup> <https://www.tandfonline.com/eprint/G6RS4EWfR8H6iPsFwbQ6/full>

<sup>5</sup> [https://peasec.de/wp-content/uploads/2019/01/2019\\_ReuterHaKiSc\\_FakeNewsPerceptionGermany\\_WI.pdf](https://peasec.de/wp-content/uploads/2019/01/2019_ReuterHaKiSc_FakeNewsPerceptionGermany_WI.pdf)

<sup>6</sup> [https://www.ted.com/talks/eli\\_pariser\\_beware\\_online\\_filter\\_bubbles?referrer=playlist-how\\_to\\_pop\\_our\\_filter\\_bubbles](https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles?referrer=playlist-how_to_pop_our_filter_bubbles)

media and information literacy—to identify and detect manipulated content, either as text, images, videos, or audio, are needed to counter manipulation attempts by different actors.”<sup>7</sup>

Technical solutions for publishers could include online tools and other resources. For example, The Fraunhofer Institute for Digital Media Technology (IDMT)<sup>8</sup> offers scientific methods and strategies for content analysis projects, such as audio forensics to detect and code traces within audio files or streams. The IDMT team also evaluates artificial intelligence tools and privacy-by-design setups. The Fraunhofer Institute for Secure Information Technology (SIT) offers additional resources for publishers to check the authenticity of several kinds of misinformation such as image and video manipulations, as well as wrong or misleading content published with a manipulative intent.<sup>9</sup> The SIT project DORIAN is focused on semi- and fully-automated fake news detection.

## CHECKLISTS AND RESOURCES

Every newsroom and editorial team should have a set of tools to spot and combat fake news:

- [Pre-equipped toolkits](#) with several features are a good start<sup>10</sup>
- [Browser plugins](#) have proven to be both efficient and helpful<sup>11</sup>
- For quick daily checks, [the fake news detector](#) is an easy-to-use platform<sup>12</sup>
- The International Federation of Library Association’s [infographic shows eight steps to spot fake news](#),<sup>13</sup> based on factcheck.org’s strategy for [identifying fake news](#)<sup>14</sup>
- Cornell University’s computer science department conducted a study on their fact-checking system and shows how to [automate fact-checking routines](#) in newsrooms<sup>15</sup>

## CONCLUSION

“On an average day, people around the world come into contact with hundreds of algorithms embedded into the software that operates communications, utilities and transport infrastructure, and powers all kinds of digital devices used for work, play and consumption. These algorithms have a disruptive and transformative effect, reconfiguring how systems operate, enacting new forms of algorithmic governance and enabling new forms of capital accumulation,” says Rob Kitchin,<sup>16</sup> who urges us to be acutely aware of such effects.

Kristen E. Martin at George Washington University offers a different, wide-ranging perspective: “Algorithms are an important actor in ethical decisions and influence the delegation of roles and responsibilities within these decisions. As such, firms should be responsible for not only the value-laden-ness of an algorithm but also for designing who-does-what within the algorithmic decision. As such, firms developing algorithms are accountable for designing how large of a role individuals will be permitted to take in the subsequent algorithmic decision.”<sup>17</sup>

Entrepreneur Michael Queralt points out, “algorithms are being deployed in mass in an effort to streamline and improve certain process, making obfuscated decisions that have a long term impact on the same individuals which data was compromised in the

---

<sup>7</sup> <https://oxfordre.com/communication/view/10.1093/acrefore/9780190228613.001.0001/acrefore-9780190228613-e-809>

<sup>8</sup> [https://www.idmt.fraunhofer.de/content/dam/idmt/documents/IL/av-analysis-idmt-2019\\_web.pdf](https://www.idmt.fraunhofer.de/content/dam/idmt/documents/IL/av-analysis-idmt-2019_web.pdf)

<sup>9</sup> <https://www.sit.fraunhofer.de/de/itforensics/analyse-und-verbesserung-it-forensischer-werkzeuge/>

<sup>10</sup> [https://medium.com/@simon\\_black/a-toolkit-for-spotting-fake-news-59dd211c57d4](https://medium.com/@simon_black/a-toolkit-for-spotting-fake-news-59dd211c57d4)

<sup>11</sup> <https://ndsfieldnotes.wordpress.com/2019/04/16/fake-news-browser-extensions/>

<sup>12</sup> <http://www.fakenewsai.com>

<sup>13</sup> <https://www.ifla.org/publications/node/11174>

<sup>14</sup> <https://www.factcheck.org/2016/11/how-to-spot-fake-news/>

<sup>15</sup> <https://arxiv.org/abs/1904.02037>

<sup>16</sup> <http://futuredata.stanford.edu/classes/cs345s/handouts/kitchin.pdf>

<sup>17</sup> [https://www.researchgate.net/publication/324896361\\_Ethical\\_Implications\\_and\\_Accountability\\_of\\_Algorithms](https://www.researchgate.net/publication/324896361_Ethical_Implications_and_Accountability_of_Algorithms)

first place,” which is why he encourages us to invest in technology education as “data mastery is no longer an afterthought, but a core component of the organizations and its impact will be felt for years to come.”<sup>18</sup>

Algorithms are here to stay. Almost all the research cited in this paper points in the same direction: the usage of algorithms expands with consequences that have yet to be understood. This is great because it creates new jobs and new opportunities for media and journalism as well as for readers. But it is also dangerous because the emergence of algorithms is a one-way street without the possibility for turning back. We must cope with the challenge as best as we can by educating our staff and readers. Research can help and so, too, can partnerships with technology companies, universities and political institutions. Good algorithms lead to good results and better online experiences. Bad algorithms lead to filter bubbles, misinformation and fake news. Knowing how, when and why this happens is the fundamental first step of the journey.

--

This Hotsheet is presented as part of the Age of Disinformation project.



We acknowledge the support of the Government of Canada.



---

<sup>18</sup> <https://medium.com/dataseries/the-impact-of-data-algorithms-the-law-of-unintended-consequences-2b580c287661>