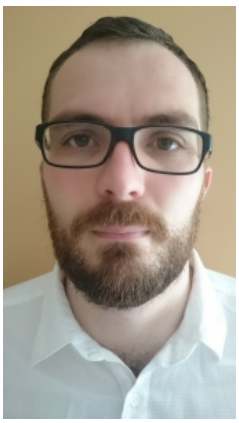
The hidden human labour behind search engine algorithms



Everybody knows that search engines use algorithms, but few know how these work and who builds them. ***Paško Bilić,*** *Research Associate at the Institute for Development and International Relations in Zagreb, Croatia, writes here about the layers of human labour behind Google's algorithms and their implications for search neutrality. His post is based on a paper published in Big Data & Society titled "Search algorithms, hidden labour and information control."*

Billions of people worldwide use search engines without understanding the inner workings and processes providing the information that they are looking for. Despite widespread use, the background mechanics of search are not fully visible and are often hidden behind technical decisions and trade secrets. Web search is a highly profitable business, from which Google secured 74.9 billion US dollars of revenue in 2015. Around 90 percent of that revenue comes from advertising.

Google's search was initially based on the PageRank algorithm, which ranked websites in terms of the number of links they received from other websites. Over time, the number of signals for ranking websites has increased significantly. Global use of the search engine, increasing popularity of social media, mobile phones, tablets, and the inclusion of new search relevance signals such as previous search history, location, terms on websitesand freshness of content, all contribute to the increasing complexity of the algorithms.

While promoted as technically neutral, the algorithms create fragmentation as search results are tied to different locations enabling bigger profit from local advertisers and online marketers. Being first in search engine results has high value for internet users seeking information and organisations seeking public exposure and visibility. Local online advertising investments flow directly to Google in a global economy of scale.

Despite increasing public awareness of these processes, there continues to be a need for better understanding of how broadly algorithms are embedded in the digital economy and culture, and of the various forms of digital work and labour that sustain and promote their utility and relevance.

The algorithmic production process requires work from highly trained professionals, such as top-level computer scientists and software engineers. Their work also becomes entangled with the input of low-skilled workers necessary for the customisation, testing, adaptation and sustainability of search in local markets worldwide.

At least three types of search labour can be identified in Google's global production process. First, the paid work of in-company engineers. This is the most celebrated aspect of algorithmic production and a major part of the story of Google's technical superiority. Second is the non-paid labour of internet users, which contributes to value generation of the company. Simply put, the more people use the search engine, the more data the company collects, analyses, packages, and ultimately sells to advertisers. Third is the least transparent and discussed labour performed by "search quality raters" and "precision evaluators" hired via third party companies specialised in crowdsourcing global workforces. Google performs so-called precision evaluations and quality assessments of algorithmic changes on a regular basis. According to the latest data published by the company, it performs around 40,000 evaluations a year.

Companies such as Lionbridge Technologies Inc. offer specific services such as enterprise crowdsourcing and search relevance testing for major tech-companies interested in expanding their businesses in local markets. They provide new outsourcing options for global enterprises

such as Google by utilizing digital technologies and distributed crowdsourced workers. According to the annual market report for Lionbridge, Google's business accounted for 11% of their 560 million dollar revenue in 2015. Google does not fully disclose details of this collaboration.

However, it does explain how precision evaluations are performed and what types of tasks are required. The most important requirement is that the human "quality raters" represent the users in the locale in which they evaluate search results. Some of the most important page quality rating tasks include considerations of content quality, website information, reputation, expertise, authoritativeness and trustworthiness. The only skill requirements for raters to have is that they demonstrate web using skills in their locale, that they understand Google's guidelines and practice implementing those guidelines.

Unpacking the layers of labour processes behind algorithmic production has wide scholarly implications and potential policy impact in the field of digital literacy, market competition, transparent, fair, innovative and open digital economy. In the past few years, multiple European Union anti-trust investigations targeted three key areas: Google's comparison shopping service, pre-installation of Google's applications and services on Android OS, and restriction of third-party websites from displaying search ads from Google's competitors.

Google built its public image on separating the so-called organic, or "untampered" search results, from paid search results or ads. This division becomes less visible from a labour perspective. Tracing and mapping the input of precision evaluators delivered for algorithmic calculations of highly skilled engineers is hidden behind a wall of trade secrets. Thus the algorithmic ranking problem ceases to be a technical problem. Transparency of algorithmic calculations is also a political and economic issue since it affects visibility of actors seeking public exposure and impacts the advertising revenue flow.

A key issue is untangling how algorithms are produced and what the role of human labour is in technically mediated digital environments. The use of human quality-raters points to an increasing awareness by Google about multiple points of departure for interpreting search results in different locales. Human culture and reflexivity are proving difficult to sort out and rank under a mathematical function, regardless of the complexity of the function. This leads to a clash between the techno-utopian vision of Google's search, the cultural nuances of different locales, and regulatory demands. There is a lot at stake for all actors whose digital existence relies on Google. When the technical vision promoted by the company reaches a point when it stops being convincing, the company loses its aura and puts its global dominance in jeopardy.

This post gives the views of the author and does not represent the position of the LSE Media Policy Project blog, nor of the London School of Economics and Political Science.

February 21st, 2017 | Algorithmic Accountability, Featured, Filtering and Censorship | 0 Comments

---