

Making and Measuring News: Data and Algorithms in Journalism

 blogs.lse.ac.uk/polis/2016/02/24/making-and-measuring-news-data-and-algorithms-in-journalism/

2016-2-24

By Alison Powell, Assistant Professor and Programme Director of the MSc in Media and Comms (Data and Society) at LSE

Journalists have a special responsibility in democratic society. They are supposed to source and report information that is of public value, investigate issues of broad public interest, and check facts.

This means that journalism has always been about telling stories, but has also always been about data: the data underpinning the facts, and now, new data about their audiences. But how should journalists approach data – either within the newsroom or as way of engaging their audience – and how should they engage with algorithmic systems to search, manage or analyse that data? Do data practices and use of algorithms change the public function of journalism?

On 24th February, LSE Data and Society and Polis are convening a workshop to discuss these questions, with participation from a group of academics, innovators and data journalists. We will examine four key shifts that influence the work – and the political position – of journalists. First, we'll look at the use of data in driving the production of stories, and the role of algorithms and numbers in producing these stories. Then we'll look at how data and algorithms are used within the newsroom, especially how audience metrics and analytics shift the practices of producing news, as journalists now anticipate audience responses. Yet this challenges existing work practices and raises questions about what kinds of democracy they imply.

Following the event, a report will identify the key areas for future consideration by journalists and researchers.

Data Driven Stories

Data is now a source of story ideas. The expansion of data journalism has focused on the use of data, facts and statistics as compelling material for stories and the use of computation to find stories. Open data is expanding as governments publish public data, which journalists can often gain access to, potentially tapping into new sites for public interest discussions. Web-based platforms provide ways of publishing data-driven stories so that audiences can dig down into the details, and social media distribution makes more stories accessible.

Using algorithms, open and closed data to source stories creates new business models, but also changes the way journalism is done. Investigative journalists now discover stories in data dumps and automation in newsrooms allows for processing and presentation of data that would not have been used previously. Open data sets allow audiences to hold journalists to account.

Access to Data and Trust in Algorithm

However, not all data remains open – social media platforms are now tending towards enclosure of data. Is it possible to develop more open data processes like the ones that the Guardian Data Blog journalists used to share the data that they used in their investigations by publishing the datasets associated with their articles and maintaining a dataset catalogue on their website? If data are to help identify topics for stories or material to construct them, how can all areas of the world be equally well represented?

Social media data can also be mapped and mined in order to derive news patterns leading to new journalistic business models as well as ways of understanding how to define what to cover and how to cover it. The use of algorithms for sourcing stories may be a new way of generating patterns and insights from data, but it may also raise

significant ethical challenges. Which algorithms are to be used, and what is their function? In marketing, machine learning systems' 'false positives' are not a problem, but for journalists such naturally occurring errors may lead to serious misinterpretations of publicly valuable information.

Analytics and Engagement

Audiences are now measured and audience engagement matters – as journalism has moved from assuming that journalists should not know about their audiences to believing that they must. Chris W Anderson of CUNY argues that this has some consequences for democracy: adherence to business-oriented outcome frameworks create a more market-driven role for journalists.

However, other research highlights the democratic virtues of numbers and algorithms, including the 'computer-assisted' tradition born in the US in the 1960s, which used methods from social science to quantify social issues such as inequality, and the use of rankings to present transparent information to permit better-informed individual decisions.

There is thus a key debate about whether data-based engagement shifts journalism too much towards business logics. With the increasing use of metrics, editorial decisions might be moving towards anticipating or shaping particular forms of audience engagement. Will this mean that potentially viral stories will win out in the editorial discussion over more mundane ones? Will analytics always increase filter bubbles? What role do emotions play in constructing stories and judging their impact?

Our event will provide an opportunity to delve deeply into these questions – look out for our report coming soon.

@a_b_powell

- Copyright © 2014 London School of Economics and Political Science