52. Data Journalism With Impact

Paul Bradshaw

Abstract

Data journalism with impact: How and why impact is measured, how that has changed, and the factors shaping impact.

Keywords: impact, engagement, data journalism, analytics, investigative journalism, data quality

If you have not seen *Spotlight* (2015), the film about *The Boston Globe*'s investigation into institutional silence over child abuse, then you should watch it right now. More to the point—you should watch right through to the title cards at the end.¹

A list scrolls down the screen. It details the dozens and dozens of places where abuse scandals have been uncovered since the events of the film, from Akute, Nigeria, to Wollongong, Australia. But the title cards also cause us to pause in our celebrations: One of the key figures involved in the scandal, it says, was reassigned to “one of the highest ranking Roman Catholic churches in the world.”

This is the challenge of impact in data journalism: Is raising awareness of a problem “impact”? Does the story have to result in penalty or reward? Visible policy change? How important is impact? And to whom?

These last two questions are worth tackling first. Traditionally, impact has been important for two main reasons: Commercial and cultural. Commercially, measures of impact such as brand awareness and high audience figures can contribute directly to a publication's profit margin through advertising (increasing both price and volume) and subscription/copy sales (Rusbridger, 2018). Culturally, however, stories with impact have also given

¹ <https://www.imdb.com/title/tt1895587/quotes/qt3112625?mavIsAdult=false&mavCanonicalUrl=https%3A%2F%2Fwww.imdb.com%2Ftitle%2Ftt1895587%2Fquotes>

news organizations and individual journalists “bragging rights” among their peers. Both, as we shall see, have become more complicated.

Measurements of impact in journalism have, historically, been limited: Aggregate sales and audience figures, a limited pool of industry prizes, and the occasional audience survey were all that publishers could draw on. Now, of course, the challenge lies not only in a proliferation of metrics, but in a proliferation of business models, too, with the expansion of non-profit news provision in particular leading to an increasing emphasis on impact and discussion about how that might be measured (Schlemmer, 2016). Furthermore, the ability to measure impact on a story-by-story basis has meant it is no longer editors who are held responsible for audience impact, but journalists, too.

Measuring Impact by the Numbers

Perhaps the easiest measure of impact is sheer *reach*: Data-driven interactives like the BBC’s “7 Billion People and You: What’s Your Number?”² engaged millions of readers in a topical story; while at one point in 2012 Nate Silver’s data journalism was reaching one in five visitors to *The New York Times* (Byers, 2012).

Some will sneer at such crude measures—but they are important. If journalists were once criticized for trying to impress their peers at the expense of their audience, modern journalism is at least expected to prove that it can connect with that audience. In most cases this proof is needed for advertisers, but even publicly funded universal news providers like the BBC need it, too, to demonstrate that they are meeting requirements for funding.

Engagement is reach’s more sophisticated relation, and here data journalism does well, too: At one editors’ conference for newspaper publisher Reach, for example, it was revealed that simply adding a piece of data visualization to a page can increase dwell time (the amount of time a person spends on a page) by a third. Data-driven interactivity can transform the duller of subjects: In 2015 the same company’s David Higgerson noted that more than 200,000 people put their postcodes into an interactive widget by their data team based on deprivation statistics—a far higher number, he pointed out, “than I would imagine [for] a straight-forward ‘data tells us x’ story” (Higgerson, 2015).

Engagement is particularly important to organizations who rely on advertising (rates can be increased where engagement is high), but also to

2 <https://www.bbc.com/news/world-15391515>

those for whom subscriptions, donations and events are important: These tend to be connected with engagement, too.

The expansion of non-profit funding and grants often comes with an explicit requirement to monitor or demonstrate impact which is about more than just reach. *Change* and *action*, in particular—political or legal—are often referenced. The International Consortium of Investigative Journalists (ICIJ), for example, highlight the impact of their Panama Papers investigation in the fact that it resulted in “at least 150 inquiries, audits or investigations . . . in 79 countries,” alongside the more traditional metric of almost 20 awards, including the Pulitzer Prize (Fitzgibbon & Díaz-Struck, 2016; “ICIJ’s Awards,” n.d.). In the United Kingdom, a special place is reserved in data journalism history for the MPs’ expenses scandal. This not only saw *The Telegraph* newspaper leading the news agenda for weeks, but also led to the formation of a new body: The Independent Parliamentary Standards Authority (IPSA). The body now publishes open data on politicians’ expense claims, allowing them to be better held to account and leading to further data journalism.

But policy can be much broader than politics. The lending policies of banks affect millions of people, and were famously held to account in the late 1980s in the US by Bill Dedman in his Pulitzer Prize-winning “Colour of Money” series of articles. In identifying racially divided loan practices (“redlining”), the data-driven investigation also led to political, financial and legal change, with probes, new financing, lawsuits and the passing of new laws among the follow-ups.³

Fast-forward 30 years and you can see a very modern version of this approach: *ProPublica*’s “Machine Bias” series shines a light on algorithmic accountability, while the Bureau Local tapped into its network to crowd-source information on algorithmically targeted “dark ads” on social media (McClenaghan, 2017). Both have helped contribute to change in a number of Facebook’s policies, while *ProPublica*’s methods were adopted by a fair housing group in establishing the basis for a lawsuit against the social network (Angwin & Tobin, 2018; “Improving Enforcement and Promoting Diversity,” 2017; Jalonick, 2017). As the policies of algorithms become increasingly powerful in our lives—from influencing the allocation of police, to Uber pricing in non-White areas—holding these to account is becoming as important as holding more traditional political forms of power to account, too (Chammah, 2016; Stark, 2016).

What is notable about some of these examples is that their impact relies upon—and is partly demonstrated by—collaboration with others. When

3 <http://powerreporting.com/color/>

the Bureau Local talk about impact, for example, they refer to the numbers of stories produced by members of its grassroots network, inspiring others to action, while the ICIJ lists the growing scale of its networks: “LuxLeaks (2014) involved more than 80 reporters in 26 countries. Swiss Leaks (2015) more than 140 reporters in 45 countries” (Cabra, 2017). The figure rises to more than 370 reporters in nearly 80 countries for the Panama Papers investigation: A hundred media organizations publishing 4,700 articles (Blau, 2016).

What is more, the data gathered and published as a result of investigations can become a source of impact itself: The Offshore Leaks database, the ICIJ points out, “is used regularly by academics, NGOs and tax agencies” (Cabra, 2017).

There is something notable about this shift from the pride of publishing to winning plaudits for acting as facilitators and organizers and database managers. As a result, collaboration has become a skill in itself: Many non-profit organizations have community or project management roles dedicated to building and maintaining relationships with contributors and partners, and journalism training increasingly reflects this shift, too.

Some of this can be traced back to the influence of early data journalism culture: Writing about the practice in Canada in 2016, Alfred Hermida and Mary Lynn Young (2017) noted “an evolving division of labor that prioritizes inter-organizational networked journalism relationships.” And the influence was recognized further in 2018 when the Reuters Institute published a book on the rise of collaborative journalism, noting that “collaboration can become a story in itself, further increasing the impact of the journalism” (Sambrook, 2018).

Changing What We Count, How We Count It and Whether We Get It Right

Advanced technical skills are not necessarily required to create a story with impact. One of the longest-running data journalism projects, the Bureau of Investigative Journalism’s “Drone Warfare” project, has been tracking US drone strikes for over five years.⁴ Its core methodology boils down to one word: Persistence.⁵ On a weekly basis Bureau reporters have turned “free text” reports into a structured data set that can be analyzed, searched and queried. That data—complemented by interviews with sources—has been

4 <https://www.thebureauinvestigates.com/projects/drone-war>

5 <https://www.thebureauinvestigates.com/explainers/our-methodology>

used by NGOs and the Bureau has submitted written evidence to the UK Parliament's Defence Committee.⁶

Counting the uncounted is a particularly important way that data journalism can make an impact—indeed, it is probably fair to say that it is data journalism's equivalent of “giving a voice to the voiceless.” “The Migrants' Files,” a project involving journalists from over 15 countries, was started after data journalists noted that there was “no usable database of people who died in their attempt to reach or stay in Europe” (The Migrants' Files, n.d.). Its impact has been to force other agencies into action: The International Organization for Migration and others now collect their own data.

Even when a government appears to be counting something, it can be worth investigating. While working with the BBC England Data Unit on an investigation into the scale of library cuts, for example, I experienced a moment of panic when I saw that a question was being asked in Parliament for data about the issue (“Libraries Lose a Quarter of Staff as Hundreds Close,” 2016). Would the response scoop the months of work we had been doing? In fact, it didn't—instead, it established that the government itself knew less than we did about the true scale of those cuts, because they hadn't undertaken the depth of investigation that we had.

And sometimes the impact lies not in the mere existence of data, but in its representation: One project by the Mexican newspaper *El Universal*, “Ausencias Ignoradas” (Ignored absences), puts a face to over 4,500 women who have gone missing in the country in a decade (Crosas Batista, 2016). The data was there, but it hadn't been broken down to a “human” level. *Libération's* “Meurtres conjugaux, des vies derrière les chiffres” does the same thing for domestic murders of women, and Ceyda Ulukaya's “Kadin Cinayetleri” project has mapped femicides in Turkey.⁷

When Data Is Bad: Impacting Data Quality

Some of my favourite projects as a data journalist have been those which highlighted, or led to the identification of, flawed or missing data. In 2016 the BBC England Data Unit looked at how many academy schools were following rules on transparency: We picked a random sample of a hundred

6 <https://publications.parliament.uk/pa/cm201314/cmselect/cmdfence/772/772vwo8.htm>

7 <https://www.liberation.fr/apps/2018/02/meurtres-conjugaux-derriere-les-chiffres/> (French language), <http://kadincinayetleri.org/> (Turkish language), http://datadrivenjournalism.net/featured_projects/kadincinayetleri.org_putting_femicide_on_the_map<

academies and checked to see if they published a register of all their governors' interests, as required by official rules. One in five academies failed to do so—and as a result the regulator Ofcom took action against those we'd identified ("Academy Schools Breach Transparency Rules," 2016). But were they serious about ensuring this would continue? Returning to the story in later years would be important in establishing whether the impact was merely short-term, or more systemic.

Sometimes the impact of a data journalism project is a by-product—only identified when the story is ready and responses are being sought. When the Bureau Local appeared to find that 18 councils in England had nothing held over in their reserves to protect against financial uncertainty, and sought a response, it turned out the data was wrong. No one noticed the incorrect data, they reported. "Not the councils that compiled the figures, nor the Ministry of Housing, Communities and Local Government, which vetted and then released [them]" (Davies, 2018). Their investigation has added to a growing campaign for local bodies to publish data more consistently, more openly and more accurately.

Impact Beyond Innovation

As data journalism has become more routine, and more integrated into ever-complex business models, its impact has shifted from the sphere of innovation to that of delivery. As data editor David Ottewell wrote of the distinction in 2018:

Innovation is getting data journalism on a front page. Delivery is getting it on the front page day after day. Innovation is building a snazzy interactive that allows readers to explore and understand an important issue. Delivery is doing that, and getting large numbers of people to actually use it; then building another one the next day, and another the day after that. (Ottewell, 2018)

Delivery is also, of course, about impact beyond our peers, beyond the "wow" factor of a striking dataviz or interactive map—on the real world. It may be immediate, obvious and measurable, or it may be slow-burning, under the radar and diffuse. Sometimes we can feel like we did not make a difference—as in the case of *The Boston Globe's* Catholic priest—but change can take time: Reporting can sow the seeds of change, with results coming years or decades later. The Bureau Local and BBC do not know if council

or schools data will be more reliable in future—but they do know that the spotlight is on both to improve.

Sometimes shining a spotlight and accepting that it is the responsibility of others to take action is all that journalism can do; sometimes it takes action itself, and campaigns for greater openness. To this data journalism adds the ability to force greater openness, or create the tools that make it possible for others to take action.

Ultimately, data journalism with impact can set the agenda. It reaches audiences that other journalism does not reach and engages them in ways that other journalism does not. It gives a voice to the voiceless and shines a light on information which would otherwise remain obscure. It holds data to account and speaks truth to its power.

Some of this impact is quantifiable, and some has been harder to measure—and any attempt to monitor impact should bear this in mind. But that does not mean that we should not try.

Works Cited

- Academy schools breach transparency rules. (2016, November 18). *BBC News*. <https://www.bbc.com/news/uk-england-37620007>
- Angwin, J., & Tobin, A. (2018, March 27). Fair housing groups sue Facebook for allowing discrimination in housing ads. *ProPublica*. <https://www.propublica.org/article/facebook-fair-housing-lawsuit-ad-discrimination>
- Blau, U. (2016, April 6). How some 370 journalists in 80 countries made the Panama Papers happen. *Nieman Reports*. <https://niemanreports.org/articles/how-some-370-journalists-in-80-countries-made-the-panama-papers-happen/>
- Byers, D. (2012, November 6). 20% of NYT visitors read 538. *Politico*. <https://www.politico.com/blogs/media/2012/11/nate-silver-draws-of-nyt-traffic-148670.html>
- Cabra, M. (2017, November 29). How ICIJ went from having no data team to being a tech-driven media organization. *ICIJ*. <https://www.icij.org/blog/2017/11/icij-went-no-data-team-tech-driven-media-organization/>
- Chammah, M. (2016, February 3). Policing the future. The Marshall Project. <https://www.themarshallproject.org/2016/02/03/policing-the-future>
- Crosas Batista, M. (2016, June 22). How one Mexican data team uncovered the story of 4,000 missing women. *Online Journalism Blog*. <https://onlinejournalismblog.com/2016/06/22/mexico-data-journalism-ausencias-ignoradas/>
- Davies, G. (2018, May 2). Inaccurate and unchecked: Problems with local council spending data. The Bureau of Investigative Journalism. <https://www.bijournalism.org/2018/05/02/inaccurate-and-unchecked-problems-with-local-council-spending-data/>

- thebureauinvestigates.com/blog/2018-05-02/inaccurate-and-unchecked-problems-with-local-council-spending-data
- Fitzgibbon, W., & Díaz-Struck, E. (2016, December 1). Panama Papers have had historic global effects—And the impacts keep coming. ICIJ. <https://www.icij.org/investigations/panama-papers/20161201-global-impact/>
- Hermida, A., & Young, M. L. (2017). Finding the data unicorn. *Digital Journalism*, 5(2), 159–176. <https://doi.org/10.1080/21670811.2016.1162663>
- Higgerson, D. (2015, October 14). How audience metrics dispel the myth that readers don't want to get involved with serious stories. David Higgerson. <https://davidhiggerson.wordpress.com/2015/10/14/how-audience-metrics-dispel-the-myth-that-readers-dont-want-to-get-involved-with-serious-stories/>
- ICIJ's awards. (n.d.). ICIJ. <https://www.icij.org/about/awards/>
- Improving enforcement and promoting diversity: Updates to ads policies and tools. (2017, February 8). About Facebook. <https://about.fb.com/news/2017/02/improving-enforcement-and-promoting-diversity-updates-to-ads-policies-and-tools/>
- Jalonick, M. C. (2017, October 27). Facebook vows more transparency over political ads. *The Seattle Times*. <https://www.seattletimes.com/business/facebook-vows-more-transparency-over-political-ads/>
- Libraries lose a quarter of staff as hundreds close. (2016, March 29). *BBC News*. <https://www.bbc.com/news/uk-england-35707956>
- McClenaghan, M. (2017, May 18). Campaigners target voters with Brexit “dark ads.” The Bureau of Investigative Journalism. <https://www.thebureauinvestigates.com/stories/2017-05-18/campaigners-target-voters-brexite-dark-ads>
- The Migrants' Files. (n.d.). <http://www.themigrantsfiles.com/>
- Ottewell, D. (2018, March 28). The evolution of data journalism. *Medium*. <https://towardsdatascience.com/the-evolution-of-data-journalism-1e4c2802bc3d>
- Rusbridger, A. (2018, August 31). Alan Rusbridger: Who broke the news? *The Guardian*. <https://www.theguardian.com/news/2018/aug/31/alan-rusbridger-who-broke-the-news>
- Sambrook, R. (Ed.). (2018). *Global teamwork: The rise of collaboration in investigative journalism*. Reuters Institute for the Study of Journalism.
- Schlemmer, C. (2016). *Speed is not everything: How news agencies use audience metrics*. Reuters Institute for the Study of Journalism. <https://reutersinstitute.politics.ox.ac.uk/our-research/speed-not-everything-how-news-agencies-use-audience-metrics>
- Stark, J. (2016, May 2). Investigating Uber surge pricing: A data journalism case study. Global Investigative Journalism Network. <https://gijn.org/2016/05/02/investigating-uber-surge-pricing-a-data-journalism-case-study/>

About the Author

Paul Bradshaw runs the MA in Data Journalism and the MA in Multi-platform and Mobile Journalism at Birmingham City University, and fits in work as a consulting data journalist with the BBC England Data Unit alongside writing books and training journalists at a range of news organizations.