

# Developing Data Journalism in the Developing World: Data, Demand, Design and Dissemination

Eva Constantaras

Internews in Kenya

Luther Plaza

Nairobi, Kenya

+254-703-588-212

[econstantaras@internews.org](mailto:econstantaras@internews.org)

## ABSTRACT

Data journalism has tremendous potential to drive transparency and uncover corruption in developing countries. Internews in Kenya developed an online and offline open data incubator system in 2013 to grow a data journalism community in Nairobi. Our approach is based on a few key principles: data, demand, design and dissemination designed to overcoming the challenges inherent in establishing data journalism in developing countries and this paper explores that process and its outcomes.

## General Terms

Measurement, Documentation, Performance, Design, Experimentation, Human Factors, Verification.

## Keywords

Data journalism, Kenya, Open data, Developing countries, Transparency, Training.

## Introduction

Data journalism has tremendous potential to drive transparency and uncover corruption in developing countries. Many donors are funding data journalism as a means to good governance and accountability. However, in countries with little or no history of independent, objective media and open data the challenges facing aspiring data journalists can seem insurmountable.

The Internews in Kenya data journalism program identified factors preventing the growth of data journalism. 1. Scant and bad data 2. A predominantly analog media without interactive web content 3. Low levels of data, digital and graphical literacy among the general population 4. Non-existent or inadequate access to information and press freedom laws can make both data access and protection from retaliation 5. Weak analytical and digital journalism capacity in newsrooms.

One of our first realizations was that there are quite a lot of steps between exposing journalists to the tools of data journalism and data journalism actually making it into local newspapers, television, radio and online. We first needed to grow an open data community and to do that we needed both an online incubator for data access and offline incubator for training and community building. After running a few data journalism boot camps to address these issues, we realized part of the problem was that unlike in the United States and Europe, our trainees didn't usually have a support network after the training. They rarely had access to data journalism desks, chapters of Hacks/Hackers or open data

civil society organizations. So when a trainee's formula produced an unexpected error, the data was too dirty, or the visualization software didn't support the local language, there wasn't anyone to turn to for help.

Without cultivating an open data community, it would be unrealistic to expect data journalism to become viable. In 2013, we designed and implemented an approach to growing a data community based on a few key principles: data, demand, design and dissemination.

## DATA DREDGER TO FILL THE DATA VOID

Our flagship site, the [Data Dredger](#), is a resource for Kenyan journalists to download, embed, and publish visualizations of Kenyan data. Our first task was to accumulate enough data for journalists to tell stories relevant to their audiences. Many exercises in the training work relied on World Bank or United Nations data for country-level analysis, but this data is hardly compelling for the average Kenyan media consumer, who is more concerned with his or her family or community than with global trends. They want to know how many mothers die prematurely in their community as compared to one county—not one country—over. Journalists can access Kenya-specific data and visualizations on Data Dredger.

We produced [Politics of Health](#) ahead of the 2013 Kenyan presidential elections to fact-check candidates' campaign promises, and visualized areas of malarial risk, maternal mortality, and anti-retroviral shortages. The Data Dredger was the only finalist from Africa in the 2013 Data Journalism Awards but more importantly, it proved to journalists that data enabled them to produce issue-driven election coverage not found anywhere else by showing, for example, that it would take 21 years to fulfill campaign promises for universal healthcare in Kenya.

Gathering the data was an exercise in community building itself. Journalists, civic hackers, academics and think tanks all collaborated in identifying, scraping, cleaning and publishing data for reuse. Now, each time we publish a topic-specific series focused on pressing health challenges such as [newborn deaths](#), [malnutrition](#) or [antiretroviral coverage](#) journalists pick up the visualizations and data and find their own story angles.

## FOSTERING A DATA DEMAND AMONG KENYAN MEDIA AND CITIZENS

With very low levels of awareness of data, increasing data literacy became a priority over the technical skills of data journalism. By taking care of the tedious process of finding, scraping, cleaning,

and verification—something that a journalist working on her own does not have the skills or incentive for—we created a shortcut for journalists, giving them a taste of open data. The visualizations started appearing in the Kenyan media and journalists came to us for help with their data story ideas.

Data journalism training activities are now equally divided between developing a data news nose and teaching tools. Common mistakes include leaving out the data source, exaggerating the findings for the sake of a dramatic headline, misrepresenting what the data is measuring and not putting the data in the appropriate context. Journalists curious about data go through a series of exercises to understand how to use data responsibly.

They critique the way data is reported in the news by going through Jonathan Stray's [Drawing Conclusions from Data](#) check list, writing stories based on Data Dredger visualizations, composing headlines for simple charts and practicing proper citation of data sources in stories. In countries where the majority of news content involves quoting politicians with opposing opinions, developing critical thinking skills and strategies for verification of data are necessary for journalists to produce quality content. These exercises also helped journalists see the value in finding news angles in data instead of simply covering the day's breaking news.

Internews as a data journalism training provider, also wanted to gain a deeper understanding of the meaning news audiences ascribe to infographics, how people understand the visualization of data and whether they retain the information shared through the visual images. In addition, as an infographics service, through [Data Dredger](#), Internews wanted to make sure, based on evidence, that the infographics provided make sense to news audiences.

The primary modes of media access are radio, television, and print newspapers. Kenyans are increasingly dissatisfied with the direction of their government and Kenya ranks in the bottom third (136 out of 175) on [Transparency International's Corruption Perception Index](#). But analysis and visual representation of government performance or development outcomes are practically non-existent in the media, which favors blow-by-blow reporting of political scandals.

Many Kenyans have not attended secondary school and therefore many were never exposed to mathematical representations such as bar charts. What we found, in a nutshell, was that bar charts are generally considered credible and "scientific," bubble charts have little to no traction with Kenyan audiences and the best form of data representation to both interest the audience and ensure they gain insight from the data are pictorial infographics. The complete study, *Audience Reception of News Infographics: Cultural and Generational Considerations* will be published later this year but a [summary of the findings](#) include the visualizations used to test for comprehension, preference and the perceived usefulness.

Above all the research study, which was limited to a survey of 49 people in Nairobi, highlights the need for a deeper understanding and tools to measure the effectiveness of data visualization in various contexts and information ecosystems.

Long-term studies to understand demand from three angles deserve more attention: the impact of data journalism capacity building activities on the content produced by trainees following the workshop, the audience's understanding of data journalism content and the impact of data journalism content on public discourse and policy making.

## DESIGNING A SUSTAINABLE DATA JOURNALISM ACTIVITY

To complement our online data portal, we developed an offline incubator for an open data community. We decided on a four-month fellowship model under which a group of talented media professionals dedicated themselves to learning new skills and producing experimental content for their home media outlets. We pitched the fellowship as a highly prestigious opportunity for journalists, graphic designers, and developers to develop specialized skills that would make their own news outlets shine. Luckily, the Kenyan media industry is huge and several outlets could afford to let go of a staff member for four months, which is not the case in a lot of developing country media markets.

For us, the fellowship served several purposes: it immersed the fellows in months of intensive training under the tutelage of data journalists, a statistician and a graphic designer from our data team. Each module built up their cumulative skills and introduced them to data that the journalists could use for real stories when they returned to their media outlets and encouraged them to work as a team to complete assignments. Most assignments were published on a [fellowship blog](#).

Through the fellowship we learned several important lessons:

- Learning data journalism skills takes time and commitment from trainees, their editors and publishers to a new way of producing news.
- Growing relationships among journalists and the people who have data (governments, think tanks and academics) is vital to making data journalism sustainable and credible in countries without a robust open data policy and portal.
- Online production and publication is vital for journalists during the learning process not only to practice new skills, but also to develop an online brand so that when they return to the newsroom, they are given the time and resources to produce data stories.
- Data journalism is teamwork and journalists, graphic designers and developers need incentives to work

together on projects (reporting grants, formal partnering agreements, awards and other forms of recognition).

- The consumer should drive decisions about what format your story takes whether it is a radio piece for a rural audience with a trusted local community radio, an interactive online visualization for a wired urban middle class or an investigative story for a print publication.

The ground-breaking investigations produced by our fellows were driven by the objective of the fellowship: stories they produce should enable citizens to advocate for better policies and make better decisions for their communities. Apps, visualizations and other interactive digital content were not, in themselves, the objective but rather tools to inform public opinion.

There was more to it than teaching tools, too. Halfway through the fellowship, we realized it was also about learning how to tell stories with data. Fellows went out to remote regions of Kenya to put faces to the data: find mothers who had lost daughters to complications of illegal abortions, men dying of preventable diseases when more county resources could have been spent on health, and people with disabilities who haven't received safety-net payments in six months. The fellows submitted complete data analysis with findings about the demographics of those most in need. They also submitted traditional feature reporting that did not seek to find out why the system was failing, even though they had the evidence in hand.

We pushed trainees to ask hard questions to explain the data—not only to look for data to confirm their hypotheses, but also to see what conclusions they could come up with by focusing on the questions raised by the data. Where had the money gone for these programs, and what policy shifts caused uneven access to medical care?

Fellow Paul Wafula, writing for the [Standard newspaper](#), exposed that for six months, the Kenyan parliament had been holding hostage the funds for a cash transfer program for the poorest Kenyans while they tried to change the allocation criteria from need to politics.

He uncovered missing funds, inefficient distribution and ghost recipients beyond the story of members of parliament fighting for a larger share of the funds to win points with their constituencies, resulting in a poor person in the richest county being eight times more likely to receive funds than a poor person in the poorest county. After the story ran on the front page for several days, the government ordered an audit to identify and remove ghost recipients, developed new vetting committees that include community leaders and enlisted a mobile banking service to distribute the funding. Several agencies funding the project raised questions about why the government allowed politicians to change the original distribution plan that had been approved by donors.

[“When the Sun Sets in Turkana; Hunger Stakes and Stripes in the North.”](#) by fellow Mercy Juma, ran as the lead news story on

January 21, 2013. The 12-minute, data-driven story is longer than any lead story anyone can remember in the history of Kenyan television. Turkana is an isolated, impoverished region of Northern Kenya long neglected by the media and government. Juma's story reveals that malnutrition in children is a growing problem in Kenya, as famine becomes more intense and frequent, and that money goes to emergency food aid, not long-term drought mitigation. Due to the massive reaction to the story from individuals and organizations—whose phone calls started flooding in before the piece had finished airing—the station established a relief fund for Turkana County within hours, as explained in the follow-up story: [“Cases of Malnutrition on the Rise in Turkana Hospitals.”](#) The fund had raised Sh1.2 million (\$14,000 USD) by the end of January. Local politicians also offered to sponsor the families featured in the stories. The family she featured was also recruited into the safety net program that was the subject of investigation of fellow Paul Wafula's story. The print version of the story [“Famine Strikes Again” \(PDF link\)](#) brought in more donations. She followed up the article with the TV story, [“Hunger Keeps Children Away from School in Turkana.”](#) Since then there have been prominent stories across the Kenyan media on the desperate famine situation.

Even more importantly for the fellows who worked on the story, the Drought Monitoring Committee asked Mercy to share footage from her story because they claimed they were not aware that the situation had become so desperate. They also requested access to her data, which she had obtained from another office also under the Ministry of Environment, Water, and Natural Resources after a long negotiation. The Committee explained that long-term drought relief strategies have been drafted but never implemented. Based on her water shortage data, the Ministry plans to travel to Turkana to dig more boreholes. The government, through the Ministry of Planning and Devolution, released 2.3 billion Sh (\$27 million USD) to go towards relief distribution in Turkana County, a development that Mercy is following closely along with the progress of the long-term drought relief legislation. NTV is eager to send her on the expensive trek to the region for follow-up stories. An overview of all the five projects produced by the fellows, links to their stories and an explanation of their impact can be found [here](#).

## **DISSEMINATING INTERACTIVE CONTENT THROUGH TRADITIONAL, DIGITAL AND SOCIAL MEDIA PLATFORMS**

The fellowship blog encouraged fellows not only to produce but also to invest in creating an online brand for themselves in a media environment where data journalism was virtually unknown. The blog built up their professional profile, disseminated Kenyan examples of data journalism through social media and helped a data journalism community coalesce. It also enriched the Data Dredger by demonstrating the fellows' evolving skills.

Establishing a brand and confidence in their work proved crucial when it came time for the fellows to return to their outlets with

their data journalism projects. Though all editors signed a contract that committed the media outlet to publishing each of the fellow's data journalism products, many were reluctant to undertake the additional work for online editors and political risk for exposing corruption. Back in the newsroom, the fellows had to fight not only for extra time to produce data stories but also for airtime, column space, and the ability to break out of their standard web templates to accommodate embedded interactive content and special project pages.

Of course, the fellowship alone couldn't transform a very traditional media apparatus and audience into data-lovers. For the first round of projects, the design of the print versions of each of the stories far outshone the online versions. In the Kenyan news cycle, page designers and television producers focus their attention and resources on traditional formats, and later an online editor uploads the content into a fixed template, often without the visuals, for a secondary audience. Our end goal was a more informed citizenry, not to win awards for interactive visuals, so for the fellows' final products, we chose easy-to-understand infographics and saved the bells and whistles for the web versions. Michael Mosota, our graphic design fellow, and Daniel Cheseret, our developer fellow, created two different versions: a version for print and an online version with more sophisticated news apps and interactive maps.

Institutionally, each of the fellows had to fight to defend their data journalism from being cut, transformed, or simplified. Mercy Juma, an Internews fellow and journalist for NTV, fought to keep infographics in her television piece and the print version that ran in the Nation (see above).

She threatened to withhold her story if the producer did not broadcast the uncut version. In order to publish a news app that enables Kenyans to explore how much their county is spending on healthcare, our developer fellow Daniel Cheseret went into The Standard newsroom and set up an independent website that was later migrated to The Standard Online after the print story generated high traffic to the app.

## **DEVELOPING A MODEL FOR GLOBAL DATA JOURNALISM GROWTH**

We have applied the same key principles of data, demand, design and dissemination to fostering open data communities in other contexts with different resources and time constraints. Unfortunately, this usually means a less thriving media industry, less money and less time. Strategies include preparing data sets in advance and hosting workshops on a specific public interest topic. We also design an incentive system for production including requiring commitment to publishing a data driven story, providing online mentoring and software prizes to reward publication and establishing cross-border partnerships with international media and open data advocacy organizations to disseminate content more widely. With Western media still mired in an industry crisis and a dearth of quality international reporting, budding data journalism movements in developing countries have an opportunity to fill that void and make data journalism truly global in both production and consumption.