The Promise and Challenge of Data Journalism

Lessons from Pakistan

Table of Contents

The Promise and Challenge of Data Journalism: Lessons from Pakistan

Contributions from Eva Constantaras, Asad Jan, Asma Usman and Muhammad Adil.

Introduction

Data journalists have a tremendous role to play as a public-service watchdog. They can transform the flood of publically available data into insight that facilitates citizen engagement in the democratic process. Yet six years after Simon Rogers, founder of the Guardian DataBlog declared: "Anyone can do it. Data journalism is the new punk," everyone is not doing it.

Despite the potential of data journalism to provide critical context, insight and actionable information in times of information crisis, most media are turning away because it is difficult, slow and expensive. Even after massive cross-border projects like the Consortium for Investigative Journalists' Panama Papers prompted political upheaval and an examination of global financial oversite, many media houses are reluctant to take the risk of investing in their own data team. If data were regarded as an asset and not a resource drain, data journalism could fulfill its essential role in the transparency and accountability cycle.

Recent research about the state of data journalism paints a dismal picture. Google News Lab's Data Journalism in 2017 reports: data skills are specialized and time-consuming to develop; time pressure prevents aspiring data journalists from producing the stories they believe need reporting; and, there are few economic incentives for media houses to invest in data teams. Data-driven reporting: An on-going (r)evolution?, a report from the University of Hamburg investigates projects that were nominated for the Data Journalism Awards, a prize issued annually by the Global Editors Network, and finds: data journalism is labor intensive; most of the coverage is of political rather than social issues; and most data is official, not collected by reporting teams.

Internews brings a commitment to accessible data journalism by, about, and for marginalized communities. By making injustice, inequality and discrimination measurable, we make it visible. If it is both measurable and visible, it may be solvable. These kinds of stories engage journalists deeply rooted in marginalized communities. To spark community conversations, the journalists tap into issues of concern to groups that face institutional discrimination. The stories they produce are disseminated through local media to reach the most people and pressure governments into making better decisions for the whole country. They prioritize storytelling over resource-intensive interactives to lower the barrier to entry and reach ordinary citizens.

Developing quality data journalism: A proven approach in Pakistan

Internews, a non-profit organization that has operated for 35 years, across more than 100 countries, works to ensure access to trusted, quality information that empowers people to have a voice in their future and to live healthy, secure, and rewarding lives.

Across the world, Internews is growing open data communities to help people to understand data, what it means, and why it matters. Internews' data journalism projects in Africa, Asia and Latin America analyze complex dynamics that cause development initiatives to succeed or fail, people to become sick or well and inequality to grow or shrink. These stories promote accountability and data-driven policy decisions in countries with little history of civic engagement.

Producing data journalism in some of the most impoverished, unequal and unsafe parts of the world for some of the most uneducated, oppressed and voiceless people has brought us to an important conclusion about data journalism. Injustice, inequality and discrimination are ubiquitous, insidious and overlooked. Learning data journalism skills is indeed labor intensive, time consuming and expensive. Yet in all countries there are journalists eager to embrace a new mindset and new tools to measure just how bad things are, who is suffering as a result, whose fault it is and how to make things better. In countries from [Kenya](#) to [Afghanistan](#), Internews trainees have launched their own data teams in mainstream media and prompted politicians to change policies and engage in a debate grounded in reality.

In Pakistan, Internews designed a robust and effective 12-month program to bring the promise of data journalism to remote tribal areas of the country. The Internews Data Journalism Certification Program was designed to produce a cadre of public-interest journalists who harness data to cover critical issues in neglected regions. The goal of the program was to catalyze the development of Pakistan's open data community and evidence-based decision making in Baluchistan, FATA and Khyber Pakhtunkhwa.

The program took place in three phases with three distinct objectives:

- **Assessment:** This phase included an open data survey, a data-driven information ecosystem assessment of media, civil society and civic tech, and a competitive selection process of participants.
- **Training:** The second phase encompassed 200-hours of classroom training program, theory and skills-based learning assessments and a simulated newsroom environment.
- **Production Support:** The final phase focused on providing structured mentoring and editorial guidance through small reporting grants, rigorous content analysis of stories and promotion and circulation of stories through Pakistani and global channels.

Assessment

Open Data Survey

Aspiring data journalists often believe that there is not enough data available in their country. This is almost never the case. Conducting an open data assessment before program launch served several purposes. First, it created a partnership with the Sustainable Development Policy Institute, a CSO committed to open data work, to create a database of all open data resources in the country. Second, understanding of data availability by sector (ie, health vs. education)

provided focus for what data journalists could reasonably report on. Finally, knowledge of available data formats helped frame the curriculum to focus on the most needed data extraction and cleaning methods.

Story Showcase

A sample of some of the best reporting from the Internews Data Journalism Certificate Program participants:

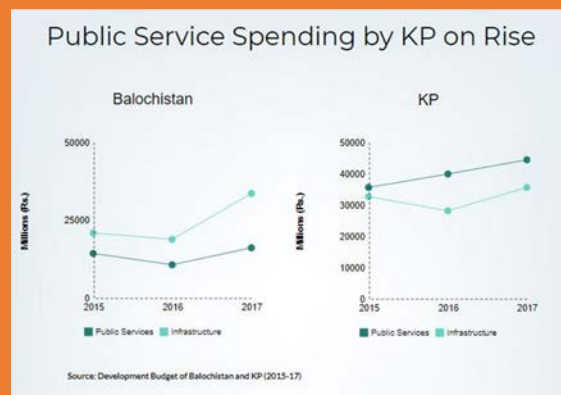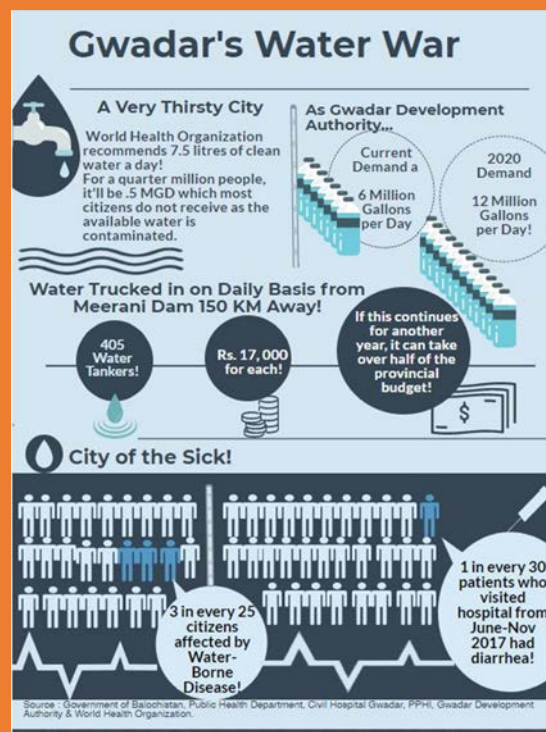*Public Health:* [Gwadar's Growing Water War](), Mariyam Suleman, PakVoices, The Balochistan Point

Summary: Booming Chinese investment and government inaction are draining Gwadar of its scant, already contaminated water resources and sent 1 in 25 citizens to the hospital for waterborne diseases in the last six months

Education and Employment: [TVET reform programmes targeting wrong skills](), Abdul Salam, News Lens Pakistan

Summary: More than half of the tens of thousands of graduates of the national TVET (Technical Education and Vocational Training) program cannot find jobs after graduation. It is not because there are no jobs to be had … just that open jobs are not the ones for which Pakistanis are being trained.

Poverty Reduction: [Balochistan Builds Roads as its Citizens Sink Further into Poverty](), Adnan Amir, Balochistan Voices

Summary: A compare and contrast story about the two poorest provinces and how divergent decisions to invest in infrastructure or public services helped Khyber Pakhtunkhwa province pull itself out of poverty and led Balochistan to sink deeper.

Information Ecosystem Assessment

This qualitative process involved interviews of all stakeholders in the data pipeline, to understand the dynamics of information in the targeted regions. Interviews with key civil society, technology, and media institutions helped identify economic, legal and other practical barriers to data-driven storytelling before the beginning of the program. Armed with this knowledge, we were able to develop alternatives for many practical challenges. By building up trust with stakeholders, we also ensured they encouraged their most promising employees to participate in the program.

Selection Process

Internews prioritized three key traits in candidates: a demonstrated ability to examine an issue from multiple angles; a commitment to telling public interest stories; and an ability to problem-solve. A total of 56 applications were received from media, civil society and the civic tech community. Candidates were screened by committee through scoring criteria designed to measure the three key traits and the Internews Data Journalism Advisor interviewed each finalists one-on-one. As a result, 14 participants were selected, with most living in target regions with a variety of journalistic, research, academic and tech backgrounds.

Capacity Building

Curriculum

Over five years working to build data teams in developing countries, Internews has developed a 200-hour in-person teaching curriculum that instills the necessary theoretical and technical skills to produce data-driven analysis and storytelling.

The pedagogy emphasizes an inquiry-based investigation. The course content balances the principles of public service journalism with the technology for data gathering, analysis and visualization. All examples and data are sourced locally, regionally, or nationally and concentrate on public interest stories that counteract corruption or mismanagement and foster governmental transparency and accountability. Course content is divided into modules, emphasizing inquiry-based investigative skills and technology labs, teaching hands-on, tool driven skills. Each lab is rooted in a public interest hypothesis that trainees prove or disprove using the skills they learn in the lab. The course devotes equal attention to data gathering and processing and to data storytelling, including effective visual and narrative techniques for data-driven communication. For a complete curriculum outline, see Annex 2.
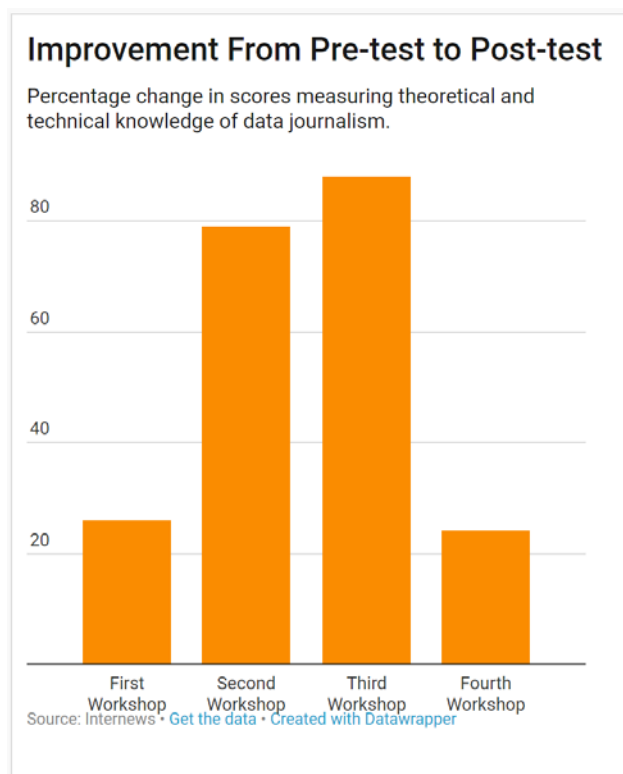
Learning Metrics

Due to the complexity and diversity of skills required for data journalism, the program closely tracked the participants' understanding of each lesson and lab. Before and after each workshop session, participants were evaluated through pre- and post-tests. The tests were designed to measure their mastery of investigative research methodology as well as technical tools. The table below presents the average percentage increase found in participants' skills and knowledge in the data journalism workshops. The content of subsequent workshop sessions was adjusted based on these results. Concepts that students struggled with were reviewed in subsequent sessions.

Student feedback during the mid-term evaluation phase included these comments:

Improvement From Pre-test to Post-test

Percentage change in scores measuring theoretical and technical knowledge of data journalism.

Source: Internews • Get the data • Created with Datawrapper

Muhammad Umer Farooq Cheema, The News: *"Prior to this training, I had cursory understanding about open data reporting. The course has improved skills regarding Excel, data cleaning and scraping. Populating data in news report is a power tool and proper persecution makes it impactful to the audience, I will try to make the best use of newly-learned skills in my future news report."*

Amna Waheed, UNDP: *"I found the step-by-step manual for Excel extremely helpful, I was not available during first week, yet were able to develop on understanding of modules conducted in first week of the program. I was already involved in open data reporting in a way, following attending this course, I see more scope for developing my skills as potential in the field. I learned many skills from this single course such as scraping data, calculation, analysis, story formulation in simplified way. I am an occasional blogger and after this training, I see a lot of scope for frequent engagements with blogging, using data journalism skills. As I am from a Civil Society Organization I can say that data skills obtained in the training shall be used to enhance analysis on personal level as well as blogger."*

The mid-term evaluation study conducted in November 2017 found that the data participants learned many new techniques and the program significantly increased the awareness and capacity of participants to utilize the latest techniques in data journalism. Some of the key findings of the evaluation are listed below:

- 80% of the participants strongly agreed or agreed that they are very satisfied or satisfied with their reporting skills after the workshops.
- 60% of the participants strongly agreed that the training helped them improved their research skills while 40% agreed that it helped improve their research skills.
- 60% of the participants strongly agreed that after the workshops, they were able to produce more data-driven stories than before, while 20% of the participants agreed that they can produce more data driven stories.
- 60 % of the participants agreed that the training was helpful in enhancing their understanding on data visualization of critical issues in the reporting. 20% agreed that that it enhanced their understanding of data visualization of the critical issues.
- 40% of the participants strongly agreed that the training developed their comprehension of what people need and want to know in news reports, while another 40% agreed with this statement.
- The most common gaps identified by many of the respondents in accessing the government records was the lack of dedicated staff to provide records in government departments. Staff are seen to treat any information as "confidential." There exists a huge communication gap between government officials and journalists.

Simulated Newsroom

The slowest, most difficult data stories are often the first stories a journalist will produce. Practice is key to the success of any sustainable data journalism capacity building program and all trainees committed to an assignment schedule in the weeks between in-person training sessions leading up to pre-workshop publication. Only students who published stories according to this schedule, which included a forensic tracking of their analysis, were admitted into the following workshop session. The following hypotheses and stories demonstrate the evolution of student's data storytelling skills as their stories grow nuanced, their datasets larger and their stories more complex.

First Assignment, October 2016
Hypothesis: Malnutrition is still a major health challenge in Pakistan.
Storyteller: Umar Cheema, Basit Khan, Muhammad Adil and Asad Zia
Story: [Number of Stunted Children Declines but Malnutrition Persists](), Data Journalism Pakistan
Summary: Pakistan's malnutrition rate is going down slowly and Pakistan spends less per capita on healthcare than other regional governments.

Hypothesis: Although the number of skilled birth attendants is growing, Pakistan's maternal mortality rate is still high.
Storyteller: Zeeshan Qasim, Mariyam Suleman & Aasia Niazi
Story: [Pakistani Mothers Die Each Day Despite the Presence of Skilled Birth Attendants in Urban Areas of Pakistan](), Data Journalism Pakistan
Summary: Maternal mortality rates are persistently high in rural areas where there is less access to antenatal and postnatal care and skilled birth attendants.

Second Assignment, December 2016

Hypothesis: Pakistan spends more than its neighbors on security at the cost of education
Storyteller: Azaz Syed
Story: [Pak. Tops In Military Spending, India On Arms Purchase And Nepal In Education](#), Data Journalism Pakistan
Summary: Pakistan spends nearly a quarter of its national budget on security, which is about what Nepal invests in education.

Hypothesis: Investing in school infrastructure has not kept girls in school in Khyber Pakhtunkhwa
Storyteller: Amna Durrani
Story: [Girls Dropping Out Of School In Khyber Pakhtunkhwa Despite Enjoying Better School Facilities Than Boys](#), Data Journalism Pakistan
Summary: Investing in girls' school facilities such as electricity and toilets is not enough to bring girls back into the classroom.

Third Assignment, January 2017

Hypothesis: The government is not investing enough in addressing Pakistan's rapidly rising rates of HIV infection
Storyteller: Aasia Niazi
Story: [Menace of HIV and AIDS in Pakistan: Ignorance in the face of rising infection rates](#), Data Journalism Pakistan
Summary: Even though Pakistan is among the minority of countries where more and more people are being infected with HIV, only half of Pakistanis know basic facts about the disease.

Hypothesis: Parents in FATA are rushing to enroll their children in private school despite poor educational outcomes in both public and private schools.
Storyteller: Abdul Salam
Story: [In KP, Parents Still Prefer Private Over Public Schools](#), News Lens Pakistan
Summary: Parents are increasingly choosing private schools even though currently, the investment per student in public and private schools is even and half of the students in both private and public schools cannot pass basic mathematics and literacy tests.

Final Assignment, May 2017

Hypothesis: Pakistan's high rate of maternal mortality could be ameliorated by equipping primary health care facilities with resources to handle birth complications.
Storyteller: Amina Amir

Story: [Maternal Health Television Clip](), [How Pakistan Treats its New Mothers](), Geo News TV

Summary: Lack of basic and primary health facilities in Pakistan, including staff, equipment and training shortages, is the leading cause behind women's deaths due to labor-related complications.

Hypothesis: The national wide vocational training program is failing because skills training does not align with skills demanded by the job market.
Storyteller: Abdul Salam
Story: [TVET Reform programmes targeting wrong skills](), News Lens Pakistan

Summary: More than half of the tens of thousands of graduates of the national TVET (Technical Education and Vocational Training) program cannot find jobs after graduation. It is not because there are no jobs to be had … just that open jobs are not the ones for which Pakistanis are being trained.

Production

Small reporting grant recipients were selected based on the strength of their pitches, which included background research, hypothesis, data and clear public interest angle. Five viable pitches were identified and funded. As during the workshop, grantees followed a story assignment schedule and their stories were mentored and edited through a Masterfile where each recipient documented their work with weekly feedback. This process ensured the soundness of the research proposal, the data integrity throughout the process, the analysis methodology and the effectiveness of the storytelling and production. (One story was not completed).

Supported Grants:

Hypothesis: Balochistan government programs for water resource management have failed to counter the water crisis in Gwadar and the health program has failed to keep up with increasing number of water-related diseases.
Storyteller: Mariyam Suleman
Story: [Gwadar's Growing Water War](), PakVoices, The Balochistan Point

Summary: Booming Chinese investment and government inaction are draining Gwadar of its scant, already contaminated water resources and sent 1 in 25 citizens to the hospital for waterborne diseases in the last six months

Hypothesis: Balochistan's growing and FATAs declining poverty can be attributed to government decisions about whether to invest in infrastructure or public services.
Storyteller: Adnan Amir
Story: [Balochistan Builds Roads as its Citizens Sink Further into Poverty](), Balochistan Voices

Summary: A compare and contrast story about the two poorest provinces and how divergent decisions to invest in infrastructure or public services helped Khyber Pakhtunkhwa province pull itself out of poverty and led Balochistan to sink deeper.

Hypothesis: Although the government of Pakistan invests to provide free services at public health centers, many opt for private healthcare despite lack of information on quality of either private or public facilities.

Storyteller: Aasia NiaziStory: [Private health care: faster and better, or just faster and available?](), News Lens Pakistan

Summary: Pakistan's population boom has overwhelmed Pakistan's stagnant public healthcare system, driving people to private institutions with shorter wait times for critical procedures, more time with physicians, but no concrete evidence of higher quality care.

Hypothesis: Lower female enrollment rates, low workforce participation and evidence of discrimination in academia reveal persistent gender discrimination even as women become more educated.

Storyteller: Abdul Salam

Story: [Education Not a Guarantee for Gender Equality in Pakistan](), News Lens Pakistan

Summary: In Pakistan, even for the few women who reach the highest pinnacles of academia, workplace discrimination holds them back despite the enlightenment and social justice promoted by universities.

Content Analysis

Internews conducted content analysis of the stories produced by program sub-recipients at the beginning, middle and end of the data journalism certification program. The aim of content analysis was to identify improvements in the quality of news as a result of Internews' Data Journalism Program. The story production process ensures the statistical accuracy of the analysis itself, a feature missing from many mainstream data journalism capacity building activities, while the content analysis evaluates how that analysis is presented. Internews used a five-point grading scale to identify how well the story met the criteria for producing high-quality, data-driven, public-interest news pieces:

1 = Very poor (Fails to meet quality standards)
2 = Poor (Below quality standards)
3 = Fair (solid performance with room for growth)
4 = Good (Above quality standards)
5 = Excellent (Meets international standards)

Additionally, qualitative comments helped Internews identify strengths and weaknesses in the news content created by the program participants as a result of the program. Internews recorded data in a spreadsheet and generated a content analysis report to identify the strengths and weaknesses of participants' published reports. 12 stories produced by six of the participants were chosen for base level analysis and 10 for advanced analysis.

The following indicators were used to analyze the stories:

1. **Lead:** The lead expresses what is newsworthy and surprising and how it affects people's lives.

2. **Data context "nut graph":** Story fully explains background of findings and data, identifies the complete source of data, and explains relevant history or technical background. The story sets data findings in context and states why new findings are significant.

3. **Accessibility: numbers expressed in ratios, rates and fractions:** Uses all simplifying techniques: rounding (rounding to a whole or easily understandable number) rates (how many people out of a total group are affected), ratios (how many people in one group are affected compared to how many in another group) and explains why the number matters. Nearly all the numbers are simplified.

4. **Data comparison among groups:** The story uses data analysis to explain the interplay among at least three datasets and their impact on the topic being covered.

5. **Evolution of the story over time:** The story includes data to explore how the issue has evolved for different groups over time and why this is a key moment for the issue.

6. **Clarity of language:** Story reads well (radio/TV: is easy to follow), contains straightforward explanations and descriptions to aid understanding, avoids ambiguous wording and jargon, and explains all specialized terms accurately.

7. **Data visualization meets all of these standards:** graph type matches data type, headline clearly states what is newsworthy about the data, source of the data is listed, colors are simple and complementary, and information from the visualization is easy to understand.

Advanced Criteria
8. **Human Interest: The story does two of the following:** The story introduces a human character, puts the person in the subject of the sentence or uses data to explain the impact on those most affected by the issue.

9. **Data analysis of issue and impact on different populations:** The story explains the consequences among different populations and their demographic characteristics.

10. **Cause analysis:** The story uses at least one data set to explain the causes of the issue.

11. **Solution analysis:** The story uses at least one set of data to explore potential solutions to the issue.

12. **Data-driven interviews:** Data findings on general trends, causes and solutions are explained by experts.
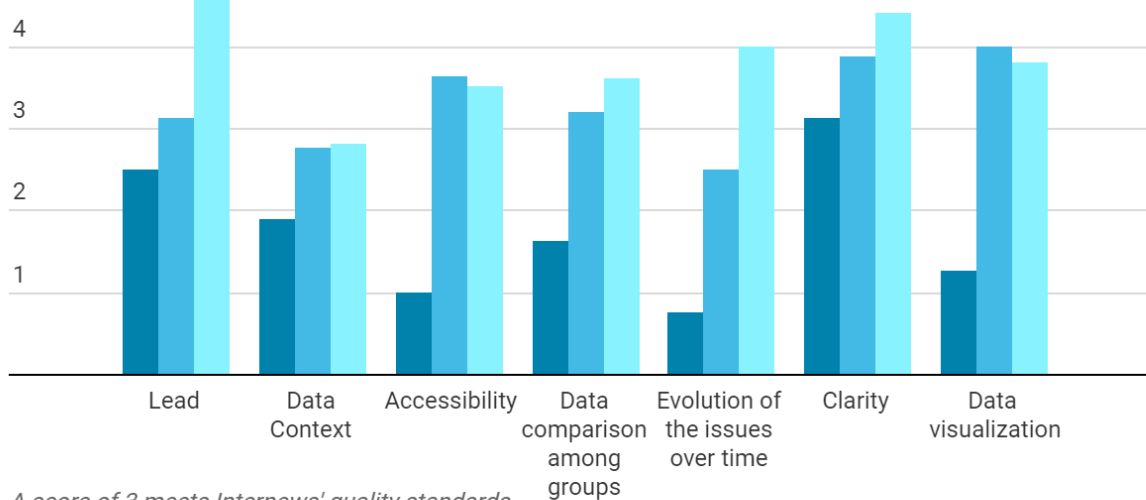
Findings

The findings of the analysis revealed that:

- The quality of written data-driven news pieces improved. Though most participants had no experience in data-driven storytelling, through the course they were able to produce effective data-driven journalism in general and most importantly, explore the root cause of development challenges through data analysis.

- Prior to attending the data journalism program, the participants were unaware of the inquiry-based research methods, but the program has significantly increased their analytical skills to harness data to explain the problem, cause, impact and potential solution.

- The pre-program evaluation of the stories indicate that many of the participants had no experience with data visualization and those who used the visualizations often did not use the appropriate graph type and lacked sources. Most of the participants mastered the use of visualizations to support their narrative and provide further detail on their topic.

- The average score overall met Internews' standards by the end of the certification program. Stories published through the grant program have yet to be evaluated.

Data Story Content Analysis Progression

Scoring criteria for basic stories are designed to evaluate potential public interest impact.

Pre-intervention End-of-intervention Post-intervention

A score of 3 meets Internews' quality standards

Source: Internews • Get the data • Created with Datawrapper

Promotion, Recognition & Impact

As a result of this intensive program, the participants produced more than 30 high-quality news pieces, which can be seen at the program's website. The site was short-listed for Global Editors' Network Data Journalism Awards for best data journalism site of the year. The Data Journalism Awards, Global Investigative Journalism Network, the European Data Journalism Conference, journalism.uk.co and the Mozilla Foundation featured stories by trainees. News Lens, Pak Voices, Balochistan Voices, The Balochistan Point And Geo News TV have all commissioned and published several data stories by trainees. Editors are expressing growing enthusiasm for data-driven content According to Google Analytics reports, data stories attract two to three times more traffic than a typical story among participating media outlets.

Balochistan Voices, the only media outlet that established a dedicated sub-site featuring data stories early enough to gather meaningful data, has seen a surge in web traffic to the sub-site. The ten data journalism stories published so far have attracted nearly 6,000 viewers in the last period, with each story averaging about eight times the number of viewer of a typical story, according to the site's most recent Google Analytics data. Currently, Balochistan Voices and News Lens are the only mainstream media outlet with their own dedicated data pages.

The stories have also caused ripples among policy makers. Following the publication of on "TVET Reform programmes targeting wrong skills," an analysis of the national vocational training program, largely funded by the German government development fund, GIZ, the government reached out to News Lens. They have since improved their five year development plans and are now working on designing demand-driven vocational training programs. The editor of News Lens also said he received a phone call from senior senator Farhat Ullah Babar who appreciated their magazine Truth Tracker and News Lens following the publication of the data stories.

The Internews Monitoring and Evaluation team conducted interviews with editors of the news organization who published data-driven news stories produced by the program participants. Through these interviews, Internews sought information on issues highlighted by the participants through their stories in Pakistan and their potential influence on different stakeholders including government. Their responses indicate that the stories attracted higher than average audience engagement and some of these stories have been picked up by the large national and international media organizations. In addition,

Following are the original quotes of the editors who directly supervise participants:

Yousaf Ajab, Editor, The Balochistan Point
"Before these stories, no data stories were ever published in our media outlet, nor any reporter were familiar about the concept of data journalism. Mariyam Suleman – Internews trainee was the first one who produce data-driven news stories for our outlet. This became one of the reasons for other reporters to take interest in data stories and showed their interest to gain knowledge as

*well. **The impact of the stories have also seen, as we recently learned that the government also took action on some of the stories related to Gwadar (including announcing a new water supply project).** We want to implement digital initiatives but we have financial constraints. We want a person who can handle data stories and video editing and production."*

Hassan Naser, Editor, Pak Voices
*"Since Pak Voices is an issue-based website, we focus on regions like Makran, Gwadar and getting good quality stories of a professional level from those areas is very difficult. Therefore, Mariyam's work was a wave of freshness and innovation for Pak Voices. **Her stories brought a substantial increase in the amount of online engagement as compared to other stories.** I believe if the top journalists also start using data in their stories, while making it user friendly with visual illustration, the public can be benefitted to a large extent. We want the reporters to be trained on the technical side on how the data stories can be managed. So, finances are required for arranging a long session of training to the citizen journalists especially from the far flung areas. And then there should be training of the trainers (TOTs) so that the knowledge can be spread and multiplied."*

Mubasher Bukhari, Editor, The News Lens
*"The data stories are the trendsetters for other reporters to follow. Following Aasia and Abdul Salam, many reporters came up with the idea of data stories. Looking at the development made after the data stories, we added a column of Data Stories in Truth Tracker and News Lens. **Our stories have been taken up by local and international media and we consider it as our achievement. The contribution to the society is that it has created an alert for the government and the policymakers that someone is working on the data and the same old rhetoric cannot go on forever.** The impact of the data stories varied; Abdul Salam started receiving threats that he will lose his job in return of being such an ambitious data journalist, also, he was constantly pressured for disclosing his source. On the contrary, senior senator Farhat Ullah Babar appreciated our magazine Truth Tracker and Newslens. Due to these data stories, almost all the provinces are working on their health and KP is constantly making an effort to introduce opportunities for their youth especially in the field of technical and vocational side. Our reports are read by legislators, policy makers and even a common person who is interested in data stories. The need of the hour is to develop the culture for data stories. Such stories need to be reader friendly, less complicated and for doing all this, we need financial support to train the reporters. Moreover, we have already developed an evaluation tool for data stories but we do not have enough finances for implementing it."*

Several trainees also won individual recognition for their work:
Name: Abdul Salam, News Lens Pakistan
Topic: Data-Driven Education Reporting
Details: This year, Abdul won two prestigious [AGAHI](#) awards for best data journalism of the year stories related to education and vocational training The stories that won him award were [TVET Reform programs targeting wrong skills](#) and [Half of FATA schools functioning in dire straits](#). The funder of the vocational training program responded directly by announcing plans to overhaul

the program, especially with how it engages women. Salam attributes his success to Internews and our trainer Eva Constantaras. In an email sent to her, he writes, "*I am dedicating these both awards to you and Internews and even further whatever achievements headed over the way waiting for me in advance dedicating you and Internews.*" Salam has also been short-listed for the Pakistan Press Foundation's Investigative Journalism Fellowship.

Name: Amina Amir, Geo News TV
Topic: Rights-based Health Reporting
Details: This year, Amina won the prestigious [CEJ-CCPP](#) Journalism Excellence Award in the category of rights-based health reporting. The story that won her this award was titled [How Pakistan treats its new mothers" in the category of Health Rights](#). The story highlights the pain and suffering pregnant women endure in Pakistan's far-flung areas due to shortage or non-availability of gynecologists, nurses and trained midwives to cater to the 13,000 babies born each day. News stories written or produced between July 1, 2016 and June 30, 2017 were considered for this award. Hundreds of entries from across the country were submitted. Amina Amir Works for Pakistan premier news channel Geo TV.

Name: Adnan Amir, Balochistan Voices
Topic: Development reporting
Details: This year, Adnan was selected for two special opportunities due to his commitment to data journalism. He traveled to London to present his experiences in data journalism at the Mozilla Foundation annual Mozfest even, bringing together media, civil society and technologists committed to a free and open internet. He was also awarded the Chevening South Asia Journalism Media Fellowship 2018 to form part of a delegation that will spend two months at University of Westminster London to grow his data journalism skills. He has created a dedicated data journalism page within his media outlet and has applied for a grant to obtain his Masters in Data Journalism in the United Kingdom.

Conclusion

Lessons Learned

The assumption that data journalism will harness data to answer basic questions about institutionalized discrimination, inequality and oppression has a major flaw. Most of these problems were invisible before and will become invisible again if journalists stop counting the causes, the victims and the wasted funds. Data journalism at its best is by, about and for those who have the least. Internews has developed a model to instill the skills that committed journalists need to harness data for public interest reporting, created a support system for ensuring story quality and developed teams that continue their important work long after the end of the program.

The barriers to data journalism identified by academics and industry are very real. Several veteran journalists with a strong history of public interest reporting lacked the basic grasp of mathematics and technology needed for the pace of the program, and were unable to complete the course. Similarly, several more junior participants were enthusiastic and excelling, but did not receive enough institutional support despite an agreement with their editors. Though several partner media outlets have embraced data journalism, they are not nationwide mass media and have limited online audiences. It remains to be seen whether after the economic and editorial support of Internews ends, if these emerging data journalists are able to carve out a niche for their new craft in the competitive Pakistani media market.

Recommendations

Data journalists have a potentially tremendous role to play as a public-service watchdog that transforms the flood of publically available data into findings that facilitate citizen engagement in the democratic process. However, considering the many barriers to widespread adoption of data journalism, it is essential that interventions are carefully designed to accommodate these realities. Programs must continuously evaluate and recalibrate activities and explore options for achieving sustainability. Several strategies would further these goals.

- Better metrics. Metrics can be applied more effectively to design (data availability, stakeholder interest, skills assessment, and audience's topical interests), capacity building (constant curriculum evaluation, rigorous real-time content analysis and statistical integrity monitoring) and post-intervention metrics (story reach and engagement, learning assessments, long-term impact tracking for policy change).
- Institutional adoption. A permanent home for data journalism capacity building, whether it be at a university, a media house or within an NGO, would vastly increase changes of long-term sustainability. Data journalism is rarely done alone and the required skill sets are constantly evolving so an in-country Center of Excellence for Data Journalism would serve the entire industry.
- Cross-border collaboration. One strategy for sustainability is to seek out a more global audience for data-driven content. Cross-border reporting on regional challenges and selling data-driven content to international media outlets are both effective strategies for reducing the financial burden on any one media house until a more effective business model for public interest reporting is found.

Annex 1: Participants

Region	Name	Organization
KPK	Asad Zia	Express Tribune
KPK	Ghulam Dastageer	The Herald
KPK	Raham Shed	Channel 24
KPK	Raza Khan	Tribal Post
KPK	Amna Waheed	UNDP
KPK	Zeeshan Qasim	Gomal University
KPK	Abdul Salam	Freelancer, Developer
Balochistan	Mariyum Mohammad Suleman	Balochistan Point
Balochistan	Muhammad Adnan	Balochistan Voices
Islamabad	Syed Ayzaz Hussain Shah	GEO News TV
Islamabad	Muhammad Umer Farooq Cheema	The News
Islamabad	Aasia Niazi	Sharp
Islamabad	Basit Ali Khan	Center for Investigative Reporting in Pakistan
Islamabad	Amina Amir	GEO News TV

Annex 2: Modules

Module 1: Introduction to Data Journalism

| Lesson 1: How Data is Used for Public Interest Stories |
| Lesson 2: How to Find a Story in Data |
| Lesson 3: Sector-Specific Data Stories |

Module 2: From Data to Stories

| Lesson 1: Common Data Formats |
| Lesson 2: Finding Data Online |
| Lesson 3: Alternative Data Sources |
| Lesson 4: Planning a Data Story: Hypothesis and Questions |
| Lesson 5: Enriching Stories With Data |
| Lesson 6: Analyzing Fact Sheets |

Labs:

Scraping and Cleaning

| Scraping 1: Scraping PDFS |
| Scraping 2: Scraping websites |
| Cleaning 1: Cleaning in Excel |
| Cleaning 2: Cleaning in OpenRefine |

Analysis

| Excel 1: Basic calculations |
| Excel 2: Summary Statistics |
| Excel 3: Percentage change |
| Excel 4: Working with large datasets |
| Excel 5: Pivot Tables |

Module 3: Understanding Data

Lesson 1: Organizing Data
Lesson 2: Verifying Data
Lesson 3: Summarizing and Simplifying Data Insights
Lesson 4:Essential Statistics
Lesson 5: Evaluating Data Interpretation
Lesson 6: Data Privacy

Module 4: Data Visualization

Lesson 1: Purpose of Data Visualization
Lesson 2:Matching Data and Graph Types
Lesson 3:Design and Color Basics
Lesson 4: Map Theory
Lesson 5: Ethics of Data Visualization
Lesson 6: Visual Storytelling

Module 5: Data-Driven Stories

Lesson 1: Reaching your Audience
Lesson 2: Data Driven Interviews
Lesson 3: Anatomy of a Data Story
Lesson 4: Data-Driven Leads
Lesson 5: Data-Driven Writing
Lesson 6: Solution Journalism

Excel 6: Advanced merging and formulas

Visualization

Visualization 1: DataWrapper
Visualization 2:Pictochart
Visualization 3 Tableau
Visualization 4: Mapping CARTO
Visualization 5:Tableau Public
Visualiation 6: Tableau Public Dashboards