Data Journalism Program Design Guidelines

April 2018

For more information, contact

Eva Constantaras, Data Journalism Trainer, Internews

@EvaConstantaras

Data Journalism Program Design Guidelines

Over the last few years, there has been a significant shift in data journalism education away from short-term activities like boot camps and hackathons. They have been replaced by more sustained and sustainable interventions that produce watchdog journalism. In the Global North and South, donors are paying attention to the importance of data driven information reaching local communities. Internews has been a leader in steering data journalism in this direction. There is a growing awareness that the challenge of teaching data journalism in many countries is split straight down the middle between teaching *data* and teaching *journalism* — in countries where neither data science nor public interest journalism are particularly well developed. Open data can be a boon to democracy — but only if there are professionals capable and motivated to transform that data into engaging and relevant information for the public.

This document includes a variety of building blocks to growing data literacy, and eventually, data journalism, in a variety of contexts.

Key Objectives

SERVE THE MARGINALIZED

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. Across the world, Internews is growing open data communities to help people to understand data, what it means, and why it matters. The Internews method teaches journalists to use investigative methods to dig deep into the institutional barriers preventing development of marginalized groups. They analyze complex dynamics that cause development initiatives to succeed or fail, people to become sick or well and inequality to grow or shrink.

REDEFINE JOURNALISM

Internews recognizes that data journalism represents a serious reformulation of the role of journalists: give citizens the information they need to make sense of the world around them and understand the forces holding back development. 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 approach includes 200-hours of classroom training program, theory and skills-based learning assessments and a simulated newsroom environment as well as structured mentoring and editorial guidance, rigorous content analysis of stories and promotion and circulation of stories through local and global channels. Our data journalists measure just how bad things are, who is suffering as a result, whose fault it is and how to make things better.

MAINSTREAM DATA

The breaking news model of journalism is dying. Learning data journalism skills is indeed labor intensive, time consuming and expensive. Yet 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. 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. Crucially, the stories they produce are disseminated through local media to reach and engage the most people and pressure governments into making better decisions for the whole country. They prioritize accessible, relatable storytelling over resource-intensive interactives to lower the barrier to entry and reach ordinary citizens.

Sample Output

Data Journalism Pakistan , Global Data Journalism Awards Finalist, 2017

Pajhwok Data, Global Data Journalism Awards Finalist, 2017

Nation Newsplex, Africa Digital Media Awards Finalist, 2016

Data Dredger, Global Data Journalism Awards Finalist, 2013

Potential Activities

Introductory Level

Open Data Survey

Internews will conduct a study to formulate an up-to-date picture of available and emerging open data resources in the country and opportunities to grow the open data community. The research, conducted over a four-week period, focuses on identifying open data resources made available by government, civil society and universities. The research process will result in a database of all existing open data platforms and data release cycles (including data published in PDF reports) in the country. Further, researchers will conduct interviews with journalists, media owners and editors, CSOs and technologists to identify the barriers and opportunities to data activities in the country.

Prerequisites: None

Resources: A local think tank with the capacity of conducting a country-wide data search or an external consultant with a track record of compiling open data sources.

Deliverable: Report on available open data resources and opportunities, database of open data in the country and a road map of challenges and opportunities faced by the emerging open data community.

Data Literacy Workshop

It is proving to be surprisingly hard to move the mindsets and approaches of even experienced journalists and activists towards an understanding of what data actually IS – before we even start helping them to DO it. Therefore, in environments with little history of civic engagement and evidence-based decision-making, a good option for **introducing open data concepts is through a series of data literacy workshops** aimed at journalists, civil society and the tech community. This activity is designed to grow the trainees' basic critical thinking and research skills to evaluate data and information for its potential value to society and how it can fit into the mission of the organization. By building awareness of the role of open data in the democratic process, we begin to engage stakeholders in a conversation about the future of open data and data-driven decision making in the country. A 40-hour general introduction to data literacy including the role of data in evidenced-based decision-making, basics of how to interpret data findings, where to find sources of data information and critical analysis of data-driven storytelling.

Day 1: Open Data, Open Government, the Data Pipeline and the Transparency Cycle

Day 2: Open Data Resources

Day 3: Interpreting Data-driven Fact Sheets

Day 4: Evaluating Data-Driven Research

Day 5: Crafting a Data-Driven Message

Prerequisites: None

Resources: An open data expert with training experience from School of Data, Open Knowledge, etc.

Deliverable: A series of workshops.

News You Can Use Data Workshop

If Internews cannot implement a full 200-hour data journalism program, this offers an 80-hour, theme-specific course for countries where open data concepts may already be better understood, or where data literacy courses have already been completed. This approach, which focuses on a specific topic and related data, can grow both data literacy and data-driven beat reporting expertise around that topic. The objective of this 10-day workshop is to develop an alternative news frame that is not preoccupied with breaking news stories of warring ideological factions and instead focuses on in-depth coverage of policy issues. It brings together a diverse community to examine whether government distribution of resources will promote development The scope of the course is very narrow, focusing on only a handful of data sets on a specific topic that a journalist could build a beat around. For example, "Using Data to Tackle the Lack of Female Participation in the Labor Market" would harness government, market, and international labor and gender data to tell short-form stories. This activity is designed to grow the trainees' basic critical thinking and research skills to evaluate a specific set of data and information for its potential value to society and how to pressure governments into basing policy in available evidence. It does NOT seek to teach the full set of skills needed for data journalism. By building awareness of the role of open data in the democratic process, we begin to engage audiences in an evaluation of the

quality of government services through both text and visuals. Popular thematic topics include gender dynamics in education, health, family planning, migration and employment.

Prerequisites: None

Resources: A trainer who has participated in Internews' 200-hour course, a local team who can identify the most relevant workshop theme and recruit appropriate participants.

Deliverable: A series of workshops and issue-specific data stories.

Data Collection Workshop

Many community-based CSOs provide services to citizens without taking advantage of rich sources of information at their fingertips to inform their own work and disseminate to decision-makers. This workshop teaches CSOs how to structure the work they do into a structured database and track recipients of services and external data relevant to their mission. This is a CSO strengthening activity that grows data literacy, data privacy practices and efficient workflows. Further, the data generated provides demographic information that can lead to better designed and more targeted programming. Often this data (for example from a service provider for domestic abuse victims or pregnant woman) can fill gaps in official data and provide valuable information to stakeholders on how to further support these organizations and communities. Finally, the data can form the core of advocacy and outreach work by these CSOs and media.

Day 1: Data collection security and integrity basics

Day 2: Building and testing data collection using toolkits such as the Open Data Kit

Day 3: Managing back-end operations of database

Day 4: Learning best practices and expertise with data management

Day 5: Data analysis and communication

Prerequisites: Data Literacy Course.

Deliverables: Databases of community organization service recipients, which can be fed into further data-driven storytelling activities.

Fact Checking Workshop: Beginner

This activity introduces data literacy skills through the lens of fact checking claims made by politicians, media and other actors. The 40-hour course focuses heavily on open data research skills, data searches, navigating data portals and evaluate the reliability of data sources to verify such claims. Ideal for election coverage, this workshop seeks to refocus the media and audience attention on policy promises made by politicians instead of on scandal and "he said" "she said" circular coverage. In the face of rising ethno-nationalist politics, this fact-checking workshop encourages journalists to take a more active role in interrogating the role of media in elections and bringing attention back to issue-based and fact-based coverage to inform citizens' vote.

Prerequisites: Open Data Survey

Resources: Local Fact-Checking expert, local staff able to attract political reporters

Deliverables: Workshops, articles fact-checking politicians on statements made in public.

Data Story Portal

Internews' data story portal makes data journalism accessible to journalists with limited digital and data literacy skills, promotes the work of emerging data journalists and the data of local CSOs and civic hackers. The data portal is a space where journalists can access data that is both raw and packaged in a user-friendly format to accommodate varying levels of data literacy. Visitors to the portal can navigate interactive infographics to access data presented in easy-to-use spreadsheets and infographics for download and publishing in their own reporting. The site also features data journalism training resources in local languages, blog posts by data journalism trainees featuring their work and sample of data driven stories that have been published and broadcast in the local media. The data journalism portal requires a fulltime team to produce sector-specific content, produce and publish training materials on an ongoing basis and coordinate contributions by external journalists.

Prerequisites: Open Data Survey

Resources: Fulltime Permanent Data Team including an editor, developer, designer and research assistant

Deliverables: Locally relevant data journalism resource site for access to data and training as well as recognition for works produced.

Intermediate Level

Fact Checking Workshop: Intermediate

The 80-hour workshop brings together journalists and CSOs with basic data literacy skills to work together on a fact –checking initiatives. This more intensive format guides teams through the process of fact checking statements using data journalism skills developed during previous activities. Trainees will learn how to group claims made by politicians by sector and how to do a deep dive into data to explain and correct policy misconceptions around that sector. For example, we may review all claims made about employment and economic growth by a candidate. This format also includes specialized CSOs and think tanks active in election monitoring to bring journalists in contact with more complex databases such as election financing, budget tracking and parliamentary monitoring.

Prerequisites: Sector-Specific Data Workshop, Data News Lab or Data Certification Program

Resources: Local Fact-Checking expert, data journalism trainer, local staff able to attract political reporters

Deliverables: Workshops, article packages fact-checking politicians on statements made in public.

Media-Civil Society Data Incubator

Journalists and civil society come together to develop data-driven projects around the data produced by a civil society organization or civic hacking group on a broad topic for 40 hours. The goal is the production of discreet, issue-driven products that aim to help society understand the root of the problem and explore solutions. For example, in a Women's Rights Media-Civil Society Data Incubator, a team of journalists from one media outlet who have participated in previous data journalism training activities are paired with a women's rights NGO that tracks cases of violence against women. In a Climate Change Media-Civil Society Incubator, a media organization is paired with an environmental protection NGO that collects data about contamination around mining sites. Together, the journalist and CSO team participate in a training program to analyze the CSO data and transform it into stories. The media house produces a series of stories around the issue. The CSO produces campaigning and advocacy materials using the same techniques. This activity is ideal for journalists with some data journalism skills who want to focus on developing a complete project and CSOs experienced in data collection interested in ensuring that data reaches citizens as well as a policy audience.

Prerequisites: News You Can Use Data Workshop or Intermediate Data Journalism Activity

Resources: An Internews data journalism trainer and a design trainer, local staff able to identify key CSOs and journalists.

Deliverables: One data-driven product each per media and CSO team around a development issue based on that CSO's local data.

"Hack Your Report" Digital Production Sprint for CSOs

Many civil society organizations, think tanks and universities are struggling to directly communicate their findings and data with their audiences. For those organizations who have gone through a data-driven communications activity, a logical next step is to take their own traditional reporting formats (commonly PDFs) and transform them into a suite of digital products. These products include short stories, visualization-driven advocacy briefs, social media infographics, GIFs, messages, and briefing materials for the media. This activity assumes participants are fairly well versed in data but need assistance in transforming their findings into information that stakeholders can act upon. Usually, this activity is confined to just two to three CSOs at once to ensure tailored content and a content production schedule that ensures the suite of products is ready by the end of the program.

Prerequisites: Introductory Data Literacy or Intermediate Data Journalism Activity

Resources: Data journalism trainer, a design trainer, local staff able to identify key CSO.

Deliverables: Suite of outreach and advocacy campaign materials based on a traditional report

Data News Lab

The fulltime six-week Data News Lab structure imitates the structure of a working data news desk. Usually, the News Lab focuses on one beat, such as economy, health care or education. It is run by a professional trainer who acts the role of Data Editor, and who is supported by (depending on the size of your budget) a deputy editor, a graphic designer, a statistician (who may also be mentoring the emerging data scientists) and a developer and a translator may be required depending on the local

context. Just like a regular newsroom, the Data News Lab is populated by trainees/participants with a variety of skills, who are allocated roles according to their level of skill and experience in journalism, graphic design and statistics. And just as in a real newsroom, participants spend a good part of their days working on the real news of the day, beginning with a newsroom meeting, decision-making about the agenda to be followed, allocation of tasks by teams, reporting, data collection and analysis activities, submissions of data and reports for editing and mentoring, copy editing, and meeting deadlines for the creation of final product (Data News Lab webpage containing print, audio, video or a mix of all the above.) To maximize sustainability, think tanks and civic technologists who gather data are integrated into the news lab to build a cooperative relationship.

Data News Lab's primary output is a Data News Lab Portal that serves as a platform for a professional network that connects and serves the media community in which it is rooted:

Prerequisites: Open Data Survey. This activity is only suitable in healthy media markets where we can secure the commitment by editors for journalists to leave the newsroom for 5-6 weeks.

Requirements: A data journalism trainer, a research assistant, a designer and a developer who can develop content for 5-6 weeks and be on the ground for 5-6 weeks.

Deliverables: Data-driven news and feature stories in local media that use data to enrich content and provide context on local issues, increased interest and commitment by editors and publishers to data journalism and trainees who produce data-driven stories in a limited time.

Fulltime or Part-time Data Journalism Certification Program

This is Internews' flagship data journalism activity and is **designed to establish data units within mainstream media houses or within universities**. Competitively selected teams of fellows from participating media outlets participate in a tiered multi-dimensional training initiative over the course of 6-18 months (fulltime to part-time). The certification program design is intended to provide them with the skills and levels of data literacy necessary to produce data-driven investigative pieces to promote accountability and evidence-based decision making. Internews collaborates with local partners and media outlets to competitively select 10-15 journalists, CSO data researchers, journalism professors and civic hackers. The training period is divided equally between training and production. Applicants must provide letters of support from their news organizations or employers committing to their full participation and attendance in the program as well as the publication of the stories produced. The program is guided by the 200-hour Internews Data Journalism Training Manual that focuses on the skills needed to produce accountability reporting including: hypothesis and question development; organization of an investigation; skills needed to acquire, analyze and understand data, documentary and human sources; long-form digital storytelling techniques; and audience engagement. Divided equally between training and production, participants will look into government public service delivery issues like healthcare, education and employment (especially for marginalized groups) that they can integrate into human-interest narratives for distribution through a variety of platforms. Teams will be trained to access, analyze and tell stories with data and produce in-depth investigative pieces that unite data analysis with powerful human-driven narrative to spark public debate about government spending and services. The participants will develop data driven multimedia stories that make development issues more transparent. They will develop a unique skillset to counter corruption and support government transparency and accountability by identifying areas where public spending does not match

up with citizens' needs and priorities and by exposing corruption and misspending. The fulltime programs run for several months with daily sessions. This option is more resource intensive as it requires on-the-ground presence of trainers for the entire period. Compressing the 200-hours into five weeks is not recommended because trainees become overloaded.

Prerequisites: Open Data Survey, commitment from editors, extensive recruiting.

Resources: Fulltime data journalism trainer and assistant trainer for training period plus mentoring grants, small grants for reporting following the program, signed letters of commitment from editors.

Deliverables: Several long-form investigative stories uncovering the root of development challenges and exploring solutions for evidence-based decision-making.

Advanced Level

Training of Trainers for Mature Data Journalism Environments

Some countries Internews works in already boasts a cadre of investigative data journalists who have played a key role in the fight against corruption in the country. Often, these journalists are self-taught and when teaching data skills, can tend to resort to teaching tools without an overarching pedagogical framework. Internews works with such journalists in a 10-day TOT to familiarize them with the Internews 200-hour data journalism training manual. It then provides a small grant or fellowship opportunity to for local trainers to localize the manual to their country through regional examples and local data. The training-of-trainers will familiarize data journalists with an approach that recognized the major challenge of teaching data-driven public interest reporting is not the tools or technology, but rather instilling a public interest mindset and applying a rigorous investigative methodology.

Prerequisites: This activity is only appropriate in the handful of environments that already have at least a dozen working data journalists.

Requirements: Data Journalism Trainer, funding for participants to localize training manual.

Deliverables: 10-day training of trainers and localized 200-hour training manual

Data Journalism Undergraduate or Graduate Curriculum

One of the most promising areas of growth has been bringing data journalism into the classroom, to redefine journalism for a new generation. If data journalism can be mainstreamed into journalism programs, it accelerates the process of "data journalism" just becoming "journalism" as championed by Simon Rogers, founder of the Guardian Datablog. However, integrating data journalism into the curriculum involves more than just tailoring the training manual. The content is designed to instill the skills and level of data literacy necessary to produce both short data stories on a deadline as well as longer in-depth system stories that explain underlying development challenges facing the region through data and digital storytelling. These constantly evolving curricula seek to balance the cultivation of critical thinking skills, research methods, coding and storytelling in a coherent specialization that produces students with the mindset and skills to integrate data into their storytelling. This requires commitment not only from the journalism faculty, but also participation from the statistics, computer science and art departments. This activity requires commitment from one or multiple universities to undergo an intensive training program for faculty that takes them through the 200-hour course content and then builds the specialization into their own media and communications department offerings.

Prerequisites: Open Data Survey, commitment from universities to a yearlong training program for faculty

Requirements: Data journalism trainer who can offer intensive training program for faculty and then lead the first year of the program while being shadowed by trained faculty.

Deliverables: undergraduate or graduate school program in data journalism.

"What Does My Government Do?" Dashboard for Media Houses

The goal of this activity is to develop beat reporting around key re-usable datasets in media houses. This activity creates a workflow for public interest data journalism that removes as many financial barriers to data journalism as possible by recycling data for beat coverage on issues citizens care about. This activity prioritizes locally relevant databases and make them accessible through an application programming interface (APIs) for regular accountability stories, documenting the analysis and visualization process for reproducibility. Examples of this include working with the civic tech community to access MPs voting, attendance and expense records for more critical coverage of politicians; liberating health data to ensure citizens are aware of quality of care; and opening education data so citizens are aware of academic rigor and inequality of the public education system. Small media houses will build their reputation for quality in-depth coverage on key issues based on these data sources.

Prerequisites: Data journalism certification program and buy-in from a media house

Resources: A fellow embedded into that media house for a minimum of one year with the capacity to manage the database and provide editorial guidance on story angles.

Deliverables: Accountability reporting portal on mainstream media site

Level Up Skills Workshops

For practicing data journalists, Internews offers advanced courses in:

- Data scraping
- SQL
- Statistical analysis with R
- Geospatial data analysis and mapping
- News App development

Prerequisites: Intermediate level data journalism team

Resources: A data journalism trainer that can be parachuted into a location to provide specialized training

Deliverables: More advanced data story production

M&E Activities
Driving Questions when designing M&E strategy

- Did the trainees tell stories important to citizens?

- Did people collaborate?

- Did the public care?

- Did we have an impact?

- Did we further government accountability?

- Did anybody get paid for his or her work?

- Did media houses decide to invest in data?

Open Data Survey

Content Analysis to measure accuracy and depth of data stories (we have a template)

Focus Groups to determine policy impact/citizen engagement

Portal for all data stories supported by the project

Sample Program Designs
Media-Oriented

Year 1	Year 2	Year 3	Year 4
Open Data Assessment	Data Journalism Certification Program or Data News Lab	Data Journalism Certification Program or Data News Lab	University Data Journalism Curriculum
Data Literacy Workshop	Media-CSO Incubator	University Data Journalism Curriculum	Good Governance Data Story Portal
News You Can Use Data Workshop	Fact Checking Workshop	Good Governance Data Story Portal	Small reporting grants
Data Story Portal	Data Story Portal	TOT for certification program alumni	What Does My Government Do? Dashboard
	Small reporting grants	Small reporting grants	

CSO-oriented

Year 1	Year 2	Year 3	Year 4
Open Data Assessment	Data Journalism Certification Program or Data News Lab	Data Journalism Certification Program or Data News Lab	"Hack your Report" Sprint

Data Literacy Workshop	Media-CSO Incubator	"Hack your Report" Sprint	What Does My Government Do? Dashboard
News You Can Use Data Workshop	Fact Checking Workshop	Media-CSO Incubator	Small Project Grant
Data Collection Workshop	Data Collection Workshop	Small Project Grant	
	Small Project Grant	Fact Checking Workshop	

Cadillac

- Full-time 4-5 person data lab team (expat resident data journalism advisor, developer, graphic designer, statistician and junior researcher)

- Activity package

  o Year 1: Open Data assessment, data literacy workshops. 200-hour data journalism training manual adapted to country context. Development of What Does My Government Do Dashboard. News You Can Use Data Workshop and data incubator run once.

  o Year 2: Part-time data journalism certification program or data news lab. Data incubator run once and "Hack your Report" run once.

  o Year 3: fulltime and part time investigative data-driven certification programs plus one news lab

  o Year 4: fulltime and part time investigative data-driven certification programs plus one news lab

  o Year 5: establishment of in-house data journalism teams and support for independent online journalism outlets plus one news lab

- Small grants to support story production

- Yearly awards ceremony for best data journalism

- Data journalism portal

- Establishment of a local Hacks/Hackers or School of Data chapter

Robust

- Resident journalism advisor plus part-time or consultant data team

- Activity Package

  o Year 1: Data literacy workshops and data news lab run once.

- o Year 2: Part-time data journalism certification program plus one news lab. Data incubator run twice.

- o Year 3: part-time data journalism certification program plus one news lab

- o Year 4: part-time data journalism certification program plus one news lab

- o Year 5: establishment of in-house data journalism teams and support for independent online journalism outlets plus one news lab

- Small grants to support story production

- Data journalism portal

- Yearly awards ceremony for best data journalism

Could make it work

- Data journalism trainer who comes in for workshops, consultant data team

- Activity Package

  - o Year 1: Data Literacy workshops, News You Can Use Data Workshop

  - o Year 2: Part-time data journalism certification program /one news lab, data incubator run twice.

  - o Year 3: part-time data journalism certification program / one data news lab. Data incubator run twice.

  - o Year 4: part-time data journalism certification program /one news lab.

  - o Year 5: establishment of in-house data journalism teams and support for independent online journalism outlets plus one news lab

- Small grants to support story production

Bare bones

- Data journalism trainer who comes in twice a year

- Activity Package

  - o Years 1-2: Data Literacy Workshops, News You Can Use Data Workshop

  - o Data incubator run as many times as possible, Data Collection workshop.

- Small grants to support story production