

Understanding Data: Can News Media Rise to the Challenge?

By Tara Susman-Peña

June 2, 2014



**National Endowment
for Democracy**

Supporting freedom around the world

The Center for International Media Assistance (CIMA), at the National Endowment for Democracy, works to strengthen the support, raise the visibility, and improve the effectiveness of independent media development throughout the world. The center provides information, builds networks, conducts research, and highlights the indispensable role independent media play in the creation and development of sustainable democracies. An important aspect of CIMA's work is to research ways to attract additional U.S. private sector interest in and support for international media development.

CIMA convenes working groups, discussions, and panels on a variety of topics in the field of media development and assistance. The center also issues reports and recommendations based on working group discussions and other investigations. These reports aim to provide policymakers, as well as donors and practitioners, with ideas for bolstering the effectiveness of media assistance.

Mark Nelson
Senior Director

Center for International Media Assistance
National Endowment for Democracy
1025 F Street, N.W., 8th Floor
Washington, DC 20004

Phone: (202) 378-9700
Fax: (202) 378-9407
Email: CIMA@ned.org
URL: <http://cima.ned.org>

About the Author



Tara Susman-Peña is senior research officer for the Internews Center for Innovation and Learning, focusing on both the evaluation of pilot projects and broader research on innovation and policy. Before joining Internews, Susman-Peña worked in audience research at NPR, where she managed the online listener panel and qualitative research initiatives. Prior to joining NPR, her years of commercial market research experience focused on branding, video ethnography, product and service development, and communication strategy in the media, technology, health, food, fashion, and automotive industries. She is the author of a 2012 CIMA report, *Making Media Development More Effective*.

Acknowledgements

This report is based on desk research and the author's interviews, conducted in person, or by Skype, phone, or e-mail between February and June 2013 with: Alberto Cairo, Ying Chan, Darcy Christ, Eva Constantaras, Sandra Crucianelli, Anyong Cui, Steve Doig, David Donald, Craig Hammer, Mark Horvit, Brant Houston, Ida Jooste, Dave Kaplan, Jennifer LaFleur, Yolanda Ma, Nils Mulvad, Dorothy Otieno, Djordje Padesjski, Miguel Paz, Aaron Presnall, Paul Radu, Margaret Renn, Giannina Segnini, Dan Sinker, Elisa Tinsley, and Jose Roberto de Toledo. Many thanks to the interviewees for their contributions.

Table of Contents

Introduction	5
Trends and Opportunities	6
Challenges	14
Meeting the Challenges	20
Conclusion	27
Appendix A: Case Studies	28
Appendix B: Resources See CIMA Website at: www.cima.ned.org	
Endnotes	35

Introduction

Worldwide, data in digital form is being produced at a dizzying pace, not only by governments, academic institutions, and private enterprises gathering it for their own uses but also as a by-product of millions of routine interactions on computers, cellphones, GPS devices, and other digital tools. Data is rapidly exploding in quantities far vaster than the capacity of civil society, commercial entities, and individual citizens to make sense of it.

This presents an opportunity—if not a duty—for news media to play an important role in helping to analyze and digest all this information through the practice of data journalism, also referred to as “computer-assisted reporting,” “data-driven journalism,” or “precision journalism.” It also presents an opportunity for the media development sector to have a notable impact on human development by equipping journalists and media institutions with the ability to make sense of data for citizens. This report considers the implications of the rise of data journalism for the media development field.

Trends and Opportunities

With openness increasingly held up as a democratic value, governments are under more pressure to open to public view the vast amounts of data they collect on their populations. A wealth of data on education, agriculture, health, finance, licensing, crime— is available, if only governments can organize it and make it available. Citizen and civil society groups are pushing for greater transparency and accountability of governments, and the call for open government goes hand in hand with a demand for open data.

The push for open government data comes both from above and from below. In June 2013 at the G8 summit, leaders signed on to an Open Data Charter, which outlines recommendations and broad best practices for releasing government data. The opening lines of the charter provide a good characterization of the increasingly global socio-political context around data and signal a growing need for journalists to act as data interpreters:

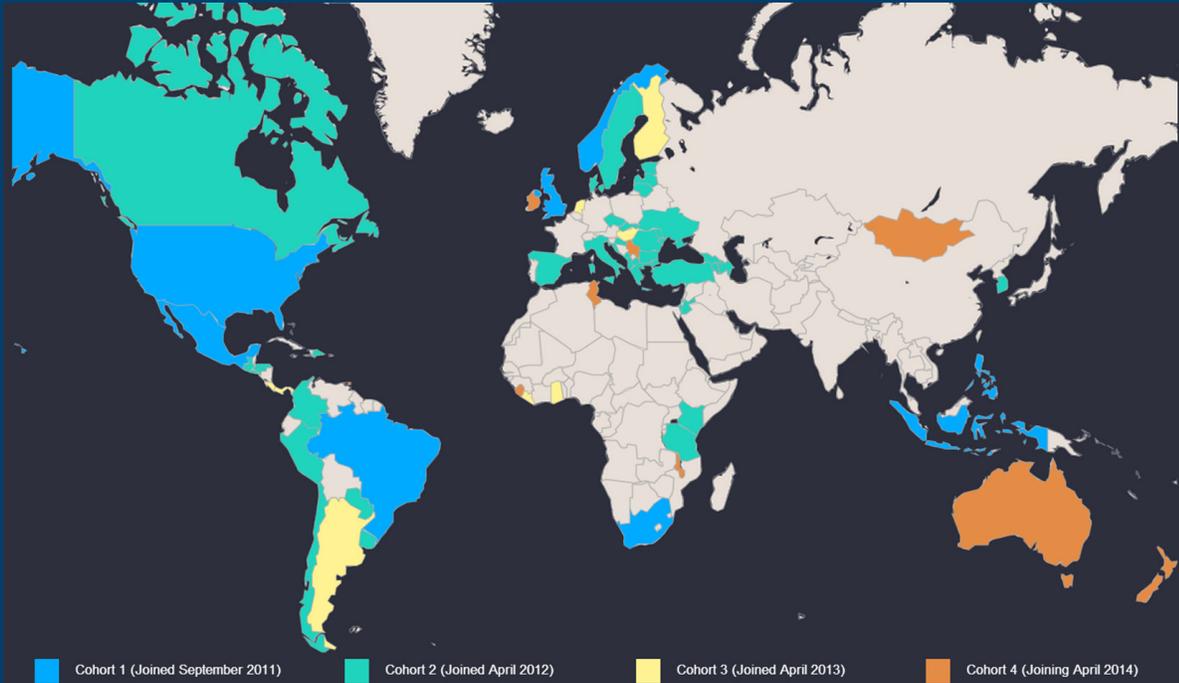
The world is witnessing the growth of a global movement facilitated by technology and social media and fuelled by information—one that contains enormous potential to create more accountable, efficient, responsive, and effective governments and businesses, and to spur economic growth. Open data sit at the heart of this global movement.¹

From civil society, the Global Open Data Initiative was launched the same month as the Open Data Charter by The World Wide Web Foundation, the Sunlight Foundation, Fundar, the Open Knowledge Foundation, and the Open Institute. The idea driving this initiative is to ensure that the open government data initiatives are useful and impactful.² The very existence of this initiative signals that simply opening government data to the public is not enough to ensure meaningful engagement with that data. The need for “infomediaries” such as data journalists to help facilitate this engagement is more salient than ever.

Under the Kenya Open Data Initiative, Kenya was the first developing country to open its government data, in 2011. All countries that sign on to the Open Government Partnership agree as part of their action plans to show progress in access to information, including facilitating access to government data.³

This is all very promising, but all countries have a way to go before journalists and others have broad access to government data—not just national data, but also data from local governments. Governments have not yet met the public demand for data that is tied to the demand for transparency and accountability. Journalism professor Steve Doig of Arizona State University observed, “This should be the default: Records should be open to the people who pay for them: taxpayers. This isn’t widespread in Europe yet.”⁴ In the countries where data has been

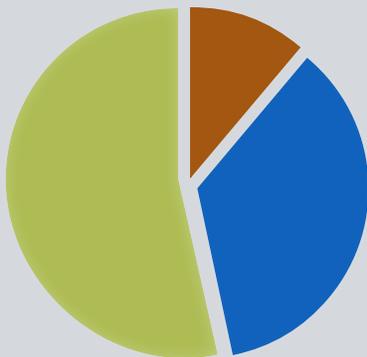
Growth in Participation in the Open Government Partnership



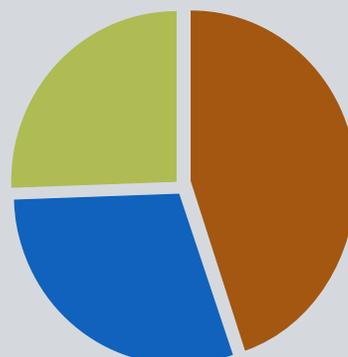
Since 2011, the number of countries participating in the Open Government Partnership has grown from 8 to the 63 indicated on the map above. Map: www.opengovpartnership.org.

Freedom House *Freedom of the Press* Ratings

Participant Countries



Non-participant Countries

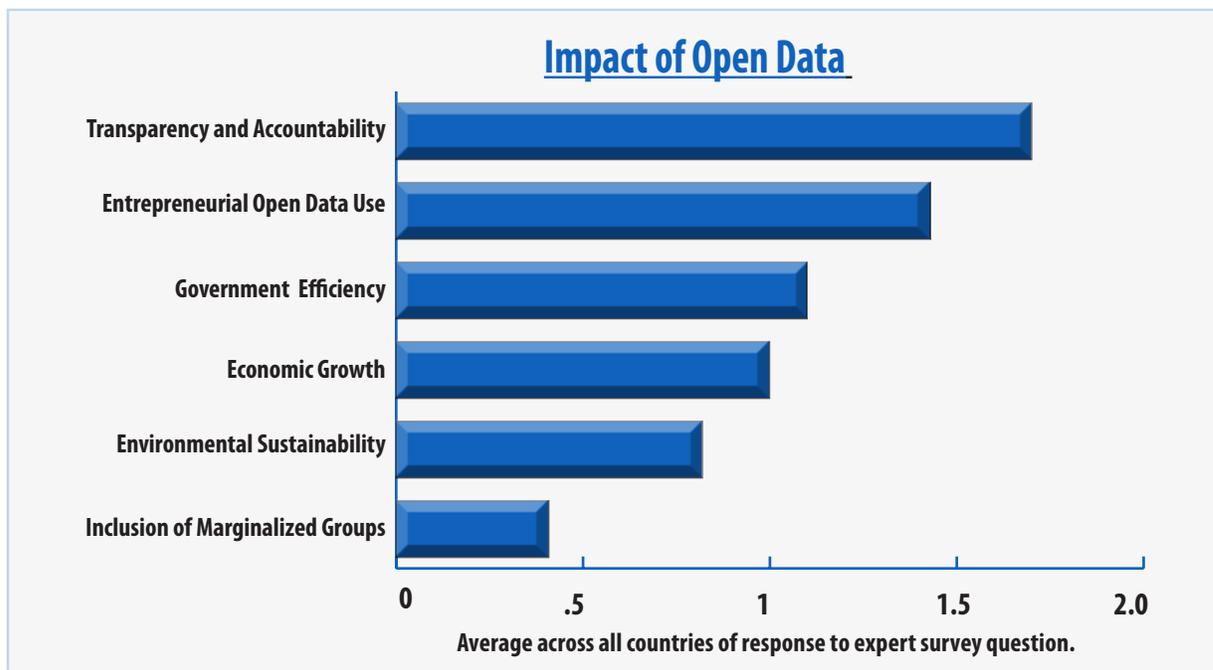


■ Free
■ Partly Free
■ Not Free

There is a substantial disparity in Freedom House ratings between countries that participate in the Open Government Partnership and those that do not.

opened, there has not been an automatic increase in government transparency or accountability. Kenya, where the legislature passed a new restrictive media law in December 2013, is an unfortunate example of a rise in countervailing forces following on the heels of an open data initiative.

Having more access to government data does not automatically increase citizen understanding or participation. The 2013 Open Data Barometer found that open data had an extremely limited impact on 77 countries in six key areas, represented below. Note that potential scores were on a scale of 0-10. While transparency and accountability scored the highest, overall the mean scores are disappointingly low.⁵



Clearly there is still a great need—perhaps now more than ever—for reliable guides through the mass of information who can pinpoint how data is relevant to people’s lives and can interpret data so that people can use it for decision-making, including participating in the governance of their countries. “People say, do we need journalists now?” said journalism professor Brant Houston of the University of Illinois. “I say, we need

them more than ever. It's one thing to do a data run; it's another to do fieldwork and interviews and analyze it."⁶ Journalists have a huge opportunity to craft stories from the government data already available and to join with advocates in continuing to press for more and better open government data.

The World Economic Forum in 2012 argued for the creation of a "data commons ... in which this information benefits society as a whole while protecting individual security and privacy."⁷ From a development perspective, the report argues, researchers with access to this data could perform analyses to more accurately determine the needs of populations and better provide services. While a global "data commons" is still a fantasy, clearly data is not just for scientists anymore.

The proliferation of data offers the opportunity for journalists to help people understand what it means and how broader trends are relevant to them. Much of the buzz around digital data has focused on new opportunities for businesses to understand consumers. But journalists, to the extent that they can access this commercial data, can be the channels through which others can learn from this information. Despite limitations in access to data in many countries, the reality, according to data journalists, is that data is growing at a pace that makes it impossible for journalists to ignore.

As various types of data have grown in quantity (more than 50 percent per year, by some estimates⁸), data analysis has become a necessary skill in a variety of professions, including journalism. As traditional journalists have grown their data skills, the flood of data has moved some non-journalists with computer skills closer to journalism, increasing opportunities for journalists and hackers to collaborate.

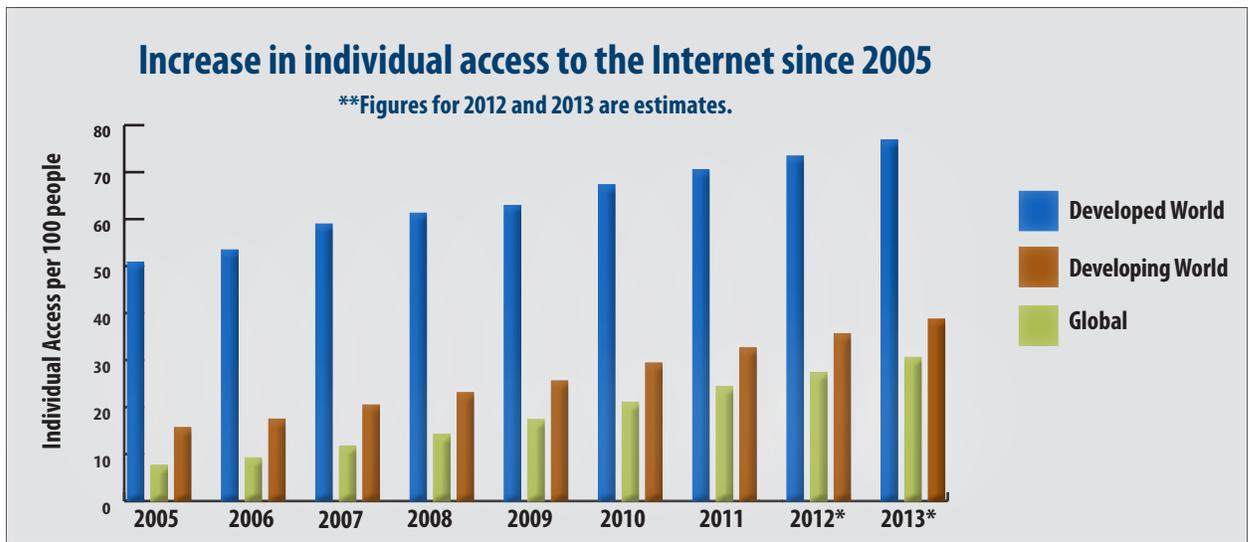
Where resources may be scarce, the Internet offers opportunities to create new types of independent media. Journalists can present stories in new ways that respond to increasing audience interest in engaging with content as opposed to simply consuming it. These trends help set the stage for reinventing the newsroom and making it more responsive to citizen needs and interests. Data reporting often leverages these trends, producing interactive applications that engage audiences.

Data journalism also benefits from broader trends in collaboration as a primary mode of work. It requires a range of skills, and is best performed by a diverse team including journalists, data analysts, designers, and programmers. Collaboration in data journalism can also enable better cross-border reporting (often leveraged for reporting on corruption), allow skill sharing, and facilitate the production of news apps. On the flip side,

cultural impediments to such collaboration can be a barrier to doing data journalism. In China, according to journalism student Cui Anyong, “The lack of communication among people of different backgrounds is the major reason that limits the development of data journalism. Everything is just at the beginning point.”

Not only has new technology produced an abundance of new data; digital technology has produced better, easier tools to work with data. Much analysis that 40 years ago required a mainframe computer that was difficult to access, complicated to operate, and costly to use, today can be done with free or low-cost Web-based tools. The open source movement, which functions by sharing and collaboration, has been a driving factor in increasing access and simplification of digital tools. Costs that can hit at various points of the analysis and reporting process—connectivity, storage, devices, the tools themselves—are dropping. The ability to visualize data analysis, from simple but powerful static bar charts to complicated interactive visualizations, has been transformed. The existence of simpler tools means that it is easier to do sophisticated analysis without having to understand as much technical detail on the programming end about how the Web-based tools function.

Underlying access to tools is a surge in connectivity—Internet access and speed. According to ITU data, the world average for individuals using the Internet is still just below 40 percent. Nonetheless, during the last decade there has been steady growth across the globe⁹:



Increased access to better and cheaper tools, of course, in no way obviates the knowledge and skills required to work with data and with the tools themselves. Data will not analyze itself. However, for those with the skills or the willingness to build them, these tools, together with rising rates of connectivity, significantly facilitate the process.

With the advent of social media and citizen journalism, the definition of journalism itself has been undergoing transformation. In the face of increased ability for citizens to publicly tell stories and engage in various of forms of reporting, some trained professional journalists' use of data gives them some distinction. "Credibility is the ultimate goal. That's why a lot of us got into this [data journalism]," Brant Houston said. The danger, of course, is that numbers can give the impression of being more trustworthy than other sources, even though data is produced by people and thus is subject to human fallibility. However, when journalists make interactive data available online, others can verify stories, check how journalists drew conclusions, and even come to their own conclusions. Some journalists have taken the opportunity to be more transparent about their newsgathering, for example offering video tutorials to explain how the story was done and guidance on using tools to further explore the data.

Not all data journalism is high-quality journalism, of course, but data journalism, when done well, has a lot to add to traditional reporting techniques. According to Ida Jooste of Internews in Kenya, data journalism "helps journalists move away from 'he said, she said,' kind of journalism; not just unverified opinion, or just political rather than factual. It helps journalists be more independent, better watchdogs, and fact-check what people are saying." Of course, data journalism is not the only way to do this kind of reporting, but it is another tool in the toolbox.

Cross border, collaborative partnerships often contribute to impactful work.¹⁰ The use of data in reporting also increases transparency in journalism, particularly when the datasets are provided for audiences to test the journalists' analysis.

A few key examples illustrate the strength and range of high quality data journalism coming from both developed and developing countries, and collaborations across regions. They are examples of projects that take on daunting amounts of data, require many months (sometimes years) of work, and employ large teams of journalists throughout the world. The scale of this work is not realistic for the large majority of news outlets, and excellent, important work can be done with much smaller, localized datasets. Nonetheless, these examples give a taste of the promise of data journalism.

Offshore Leaks project. While the International Consortium of Investigative Journalists (ICIJ) is based in Washington, DC, the mission of the project is to produce reporting about issues that cross borders using a network of journalists from around the world and a range of reporting techniques, including data analysis. The introduction to Offshore Leaks begins, “A cache of 2.5 million files has cracked open the secrets of more than 120,000 offshore companies and trusts, exposing hidden dealings of politicians, con men and the mega-rich the world over.” The reporting was intensive. “Eighty-six journalists from 46 countries used high-tech data crunching and shoe-leather reporting to sift through emails, account ledgers and other files covering nearly 30 years.”¹¹ The database for this project had to be built by the reporters for the project. Products include print (digital) articles, videos, interactive visualizations, and an eBook. In addition to the extensive reporting, the database is available for interested people to explore. There is also an explanation of how *La Nación* in Costa Rica built the database, which journalist Giannina Segnini describes with a powerful metaphor: “An art restorer is able to rebuild a broken sculpture after thoroughly studying its multiple unconnected pieces in order to decode how they related in the past, to return them to their original place and to discover its full shape again.”¹²

Connected China. Reuters’ project produced an interactive app with which users can explore connections among China’s elite leaders and map power flows across institutions. It includes interactive visualizations that draw on data and include text, photographs, video, and an accompanying text-based guide. According to Reuters, “Connected China is the culmination of 18 months of research, reporting, design and development, culminating in an app that draws on a database containing:

“Tens of thousands of entities – people, organizations, events

“More than 30,000 relationships

“1.5 million words (equivalent to more than 20 non-fiction books!)”¹³

Poderopedia, a data journalism website that maps connections across political and business elite started out in Chile and is now being developed as a network across the globe. Founder Miguel Paz realized that when “[you understand] who is connected to whom, it helps you understand why things happen.”¹⁴

Basometro, the Brazilian “Barometer” project measures support for the federal government by the congress, which represents 26 out of Brazil’s 30 political parties, and includes almost 600 politicians. The website lets Brazilian citizens explore how politicians voted, find plentiful information on politicians, and use the time slider

to see how votes evolved over time. The Basometro was created as a tool to help understand the political behavior of congressmen, but journalist Jose Roberto de Toledo was surprised when he and his colleagues at *O Estado de Sao Paulo* found that politicians were not interested in the tool and ignored it. However, the professors and researchers that Toledo feared would brutally criticize the website became its main users, and the newspaper began to collaborate with them and publish articles based on the researchers' work using the tool.¹⁵ Thus, a data visualization tool created to help clarify politics for the politicians was better used to help deepen reporting on politics.

Diaris Secretos, a hyperlocal form of reporting that was only possible using data. In the Brazilian state of Parana, a group of newspaper and television reporters teamed up to create a database from more than 700 paper diaries collected from local officials. These diaries were supposed to have been publicly available but were not. The team started with more than 200 pounds of paper, and painstakingly input the data manually into Excel, creating more than 15,000 lines of data. Their analysis of the data exposed mass corruption in the local government (e.g. the widespread use of "ghost" employees: the names of children and dead people were reported on official documents as employees). The multimedia, collaborative project revealed a multimillion dollar fraud scheme and brought down the local government.¹⁶

Challenges

The challenges to data journalism are nested within broader constraints facing journalism. These are well catalogued elsewhere, and include deficits in the areas of press freedom, a strong legal enabling environment, a business model that allows media to thrive economically, avoiding media capture by politicians and oligarchs, high quality content that reaches the population, and the ability of citizens to create, engage in, and exchange content. Even survival cannot be taken for granted; the Committee to Protect Journalists confirmed 70 journalists killed in 2013 for reasons related to their line of work.¹⁷ This means that the complications of practicing data journalism are low on the list of priorities for many. “The urge to fight for electronic records is a first world problem,” said Arizona State’s Doig.¹⁸

Another basic impediment to practicing data journalism appears paradoxical: Surprisingly, despite ever-increasing production of all types of data, journalists often have difficulty gaining access to data for stories. The challenges to access have many dimensions. Many governments are resistant to making certain data accessible to the public. Many countries do not have freedom of information laws; even with laws on the books, they are not necessarily enforced.

Making a country’s data open is a complex and nuanced process. “Even as everything from multilateral data dumps to grassroots data initiatives take off, open data evangelists discover that creating a database does not automatically lead to accessibility, participation or transparency,” observes Eva Constantaras of Internews in Kenya.¹⁹ Often, there is a dearth of government support for an open data resource. For example, local governments may lack resources to collect data, put it in an accessible format, and continually update it. Public information officers may be either unwilling or unable to answer journalists’ follow up questions about how the data was collected and what it means.²⁰ In many countries, data of interest simply isn’t collected. In the case of census and development data, this is often due to underfunded national statistics offices.

Kenya’s Open Data Initiative (ODI) was widely praised at its inception in 2011, but it has not opened up access to Kenyan data in a way that many had hoped. In fact, Fred Matiang’i, Kenya’s information and communications cabinet secretary, declared that he felt the government itself had thwarted the full potential of the ODI. In 2013 he was quoted as saying, “The fact that the Freedom of Information Act is yet to go through Parliament is a major bottleneck because there is lack of a legal backing to force institutions to release data.”²¹ Thus what is available on the ODI site is only the data that public organizations have agreed to release; they lack a specific legally obligated mandate. Because the ministries that own the data have no incentive to release data, the site is sorely lacking in

longitudinal data and granular data at the local level. Some hoped that NGOs that collect data would contribute to the site, but that process has not been implemented either. As of mid-2014, the site is a strong disappointment to those who held it up as a role model after Kenya launched its Open Government Initiative.²²

Lastly, a wealth of commercial data is available only in the private sector, either to the companies that collected it, or for purchase at a high price (and then without permission to share it publicly). Companies analyze their vast stores of consumer data to improve products and services and increase profits. While this data does not fall under the realm of a public right, it is still an area where rich sources of information could be mined for stories.

Even for the data that is available, there are no widely accepted or adhered-to standards for data. Accessible data may not be in a shape that journalists can immediately or easily use. The quality issues are many: Data may be dirty (incomplete or with mistakes or inconsistencies in how it was input), unreliable (badly collected and thus not a good representation of fact), inconsistently collected over time (some years on, some years off; or using different measurements for the same thing at different collection points), or not interoperable (different machine readable datasets cannot be integrated across different formats, systems, or organizations). Data might be recorded on paper or locked in a PDF, which makes it difficult to put into an Excel spreadsheet. Data may be presented without any available explanation of the context in which it was collected or the methodology by which it was collected, both of which impede using data as a tool for understanding.

Beyond questions of access to data, many of the challenges to more fully integrating data journalism lie in the culture of the newsroom, which has remained largely conservative despite the changes offered by digital media. Some of the challenges to data journalism are the same as the broad challenges investigative journalism faces. The 24-hour news cycle that drives newsrooms with a broadcast and/or online presence means that the practice of journalism is ongoing and fast-paced. Thus reporting that doesn't keep up with this pace is often abandoned. Breaking a story is no longer the lofty goal it once was; for countries where news is widespread on the Internet, once a story is broken in one outlet, it can quickly be picked up elsewhere. "Most reporters don't have time to do investigative reporting—to be able to look at crime numbers and make meaning out of them, to take budgets and crunch them, even for a daily beat," lamented Mark Horvit of Investigative Reporters and Editors (IRE).²³ Time is needed not only to develop individual stories, but also for data journalism to grow and mature as a profession, and for change to take hold in the newsroom. However, the intense demands of the day-to-day news business are directly in conflict with the need for more time.

Newsrooms may simply be resistant to change. The “biggest problem has to do with the internal organizational culture more than money,” according to Miguel Paz of Poderopedia in Chile. “They need convincing to foster a culture of innovation in the newsroom.”

There was a strong consensus among those interviewed for this report that data journalism does not have a hope of becoming institutionalized in newsrooms without the support of editors. Alignment of the board and executive leadership of the outlet are also crucial. But newsrooms are typically slow to change. Elisa Tinsley of the International Center for Journalists (ICFJ) explains, “Changing the overall mentality of a news organization is very challenging. Top editors often are more open to it than gatekeeper editors, because they [the gatekeepers] are overwhelmed with work. Anything new can be viewed as a burden.”²⁴ But it may not be the editors’ fault either. Very often, there is a large gap in understanding as to what data journalism is, what opportunities it offers, and what types of team members are needed to make it happen.

“Some of the media just can’t picture why you would do data journalism,” said Argentine journalist Sandra Crucianelli, “It’s not glamorous. It’s a lot of work. You are working with tables, numbers. It takes a long time to come to a conclusion, or you may not come to a conclusion at all.”²⁵

Even in newsrooms where an outlet has had success with data journalism, there may still be resistance to supporting the multidisciplinary teams needed to perform high-quality data journalism. Editors and executives often think of digital developers only in the context of support for online news infrastructure, not as potential collaborators on journalistic projects. Research revealed several instances in which a data analyst was on staff at a news outlet meant to work closely with journalists but was sitting at a desk on a different floor from the newsroom. Even when the outlet has these data professionals and developers on staff, their own training and background (not to mention interest and time constraints) may limit them from participating in journalistic activities.

At *La Nación* in Costa Rica, Segnini had been doing impactful data journalism for more than a decade when in 2009 she began to petition the paper’s board of directors for a data journalism team. She encountered resistance from the board, backed by the IT department, which claimed it would support any requests for collaboration on journalistic data projects. Segnini made requests to IT just to prove it wouldn’t work and subsequently made a presentation to the board, emphasizing the business case for a data journalism team. “It was a HUGE debate,” she said. “They didn’t want to let me have developers on the team. I fought for a year and a half. You have to be

a little crazy and really believe in it.” But in the end, her fight paid off, and *La Nación* now has a team of three journalists and two developers focusing on data journalism.²⁶

Even when the idea of multidisciplinary teams is accepted, such a structure doesn’t automatically work. Several journalists cited failures wherein international organizations attempted to “embed” developers in a news organization but didn’t create the conditions for the developers and journalists to see each other’s value and learn how to collaborate.²⁷ It is a challenging process for people of different backgrounds to learn enough of each other’s language, process, and values to work together fruitfully.

Lastly, editors and executive media leadership tend to be friendly with government officials in many countries. Often media houses are owned by politicians or the politically influential. This makes the practice of any kind of watchdog journalism challenging, and can only be addressed by attention to the enabling environment for media and to the values and culture of the news media.

Even in environments that are somewhat favorable to its practice, data journalism, like investigative journalism, will probably only ever be of interest to a fraction of the journalistic population. Journalists in countries with very low development often have low levels of education and thus are not likely to take on working with data. Journalists with some education may still be products of low-quality journalism schools. This challenge may not be immediately apparent, and in the excitement around data journalism, new initiatives sometimes get ahead of themselves. “In Kenya,” said Jooste of Internews, “people were extremely idealistic about it [data journalism]; there was a lot of hype. People thought that the data would be clean, easy to understand and compare, easily accessible. It’s been a challenge and a real concerted effort ... because of the lack of numeracy among ordinary journalists. People don’t understand what percent actually means, let alone statistical high-level stuff about turning numbers into narratives. It was really sobering to see the level of innumeracy.”²⁸

Many experts interviewed argued that journalists tend to fear math and lack even basic data literacy. (In fact, the choice of journalism as a career is often a decision to take refuge from numbers in a world of words.) Without a wide development of data skills such as statistics or basic programming, there is an extreme lack of journalists qualified for data journalism. Well-publicized errors made by neophyte data journalists only underline the need for data journalists to be trained in understanding how to verify, contextualize, and analyze data. One well-known example is a mapping error in a story uncovered by the *Los Angeles Times* that showed a massive amount of crime taking place right around the corner from

City Hall in Los Angeles. These misleading findings were due to undetected errors in the data and in the journalists' unsophisticated use of a mapping program.²⁹

Unless the media houses are willing to invest in developing these skills and interests, the practice of data journalism will neither grow nor improve. These are not skills that one comes to an end point in developing, either. Even the most skilled data journalists take on ongoing learning for themselves, keeping up their skills by training others, for example. Because of the evolving nature of the digital tools used to work with data, this is not likely to change. If journalists improve their data skills, civil society and news-engaged citizens must also improve their level of data literacy for reporting with data to have any impact.

To be fair, having enough skilled people to deal with ever-increasing production of data is a topic far beyond the concerns of just the news media. Even the United States, an early leader in data journalism and data analysis, appears to lack enough data analysts. The problem is more acute in other countries, particularly developing ones. Forget journalism; professions that owe their very existence to statistics are falling behind. As one biostatistician notes, "Our ability to accrue data is outstripping our ability to analyze it ... We need better methods, more people, faster computers. The primary bottleneck is the ability to sift through the huge amount of data that technology is producing."³⁰

With the challenge to basic data literacy and widespread analytical capacity as high as it is, it is unrealistic to expect most journalists to take on the programming that is often used in developing interactive data products. The logical course is to try to lure developers and programmers into journalism. This might be difficult on a large scale. As Dan Sinker of Knight-Mozilla OpenNews observes, there is "a lot of competition for that talent, newsrooms can't compete with the pockets of Internet companies. But with traditional journalists you could make the same argument; they could be doing PR or something else but they are journalists. ... It's an opportunity to build and experiment and do new things."³¹ There is a bit of a Catch-22: Newsrooms need to fight against the perception that all of the technological innovation is happening in the ICT sector, but without the tech muscle to help realize this innovation, it will be hard to prove.

Nonetheless, available tech tools can offer important support to the enterprise of data journalism: Many tools are free, becoming increasingly simpler to use, and improving all the time. However, challenges remain. The clearest challenge is that even the simplest tools need skilled, numerate people to operate them. Another important challenge is that most of the tools are only in English, which creates a significant and immediate obstacle for many. Many tools are

Web-based, which creates a barrier for areas with poor Internet access. Open source tools are freely available but may not provide access to technical support. And while tools are getting simpler, a programmer is still needed to do many things beyond what is available off the shelf, for example, creating interactive displays of data. A lot of the most commonly used tools are very new and a lot of open source technology that is constantly being developed and updated means that tools change and become obsolete very quickly.

Another critical question for media developers who decide to invest in data journalism is about the process by which data journalism becomes institutionalized. In countries where data journalism has become more widespread and integrated into journalistic practices, what has triggered and supported the process? There are several levels of proselytizing: from the lone journalist to her newsroom, from one newsroom to other news outlets in the area, from a local media sector across the country and the region. None of the research for this report turned up an example of a top-down change within the newsroom that instituted data journalism. Instead, change is most likely driven by one or more pioneering journalists who build buy-in from colleagues, editors, and the board of directors.

It is still too early to tell what forces support broader institutionalization from a single organization out to a sector or a country. Creating mass interest and excitement in journalism and tech communities and building demand from the public are other critical puzzles. Hackathons may be one way to create excitement around data. There may be some experiments, and a few examples of very compelling data-driven stories, which break the ice for others to follow later. In countries with open or relatively open and accessible Internet, these kinds of projects get public exposure, and space for imitators is opened.

In many countries, awareness of data journalism seems to often be sparked by one or two visionaries who have crafted an attention-getting story that puts on full display the potential for data to reveal hidden government secrets. Many interviewed for this report pointed to a handful of stories that opened up journalists' and the public's minds about the value of data journalism. The amount of Web traffic to these stories supports this perception. The tipping point toward the mainstreaming of data journalism in a country can sometimes be identified in a single story, as with the *Daily Telegraph's* 2009 exposure of British MPs' expenses abuse. The public outrage at the unfolding scandal was accompanied by a new understanding in the media of the purpose and power of using data to find stories.³²

While these challenges are significant, media developers can leverage several trends to help realize the opportunities offered by data journalism.

Meeting the Challenges

A strengthened business model

While the examples we found of the institutionalization of data journalism were journalist-driven, these journalists needed editorial and executive support to broaden the presence of data journalism in their media organizations. The mindset of media leadership and newsroom culture may be quite difficult to shift. One idea may help spark change: Data journalism may be a light at the end of the dark tunnel that the business of media entered into, particularly in the United States and Western Europe, around the time of the economic downturn of the late 2000s. Brazil, among other countries, is beginning to feel the strain, and many experts anticipated the possibility that other countries would follow this downward slope. Opinions of the experts interviewed strongly converge on this question: Data journalism should be a boon to media business. In some cases, the evidence is not yet clear that it is making additional revenue; in others, the evidence shows that data journalism is driving clicks and is thus providing advertising revenue.

There are some compelling examples of early successes of data journalism enhancing revenue. “The *Texas Tribune* and ProPublica, which are both arguably post-print media companies, reported that funding for their non-profit journalism organizations exceeded their goals much earlier than planned,” said Mirko Lorenz of Deutsche Welle.³³ A significant portion of this financial strength can be attributed to the outlets’ data journalism practice. In fact, two-thirds of the *Texas Tribune*’s traffic goes to the database pages.³⁴

One reason that data can continue to drive clicks is that it can provide a very different story or perspective to readers. The discovery and presentation of analysis from a new or unknown database; a deeper examination of an existing database; or even the creation of a database, are all unique and valuable offerings to audiences.³⁵ When a database is online, or interactive visualizations are created and displayed, the audience has an opportunity to participate in telling and understanding the stories. When a data-driven story is presented interactively, the result is an increase in clicks.

In addition to driving clicks, additional business opportunities may stem from data journalism, including selling databases. While several people interviewed mentioned this possibility, an understanding of how to offer paid databases as broadly appealing public information sources seems to be still in the beginning stages of development. There are several examples of this in the developed world, particularly around financial data. However, finance is an area that has long relied on data, so users of financial information did not have

to be converted or taught about the importance of data. Lorenz, writing in the *Data Journalism Handbook*, explains: “There are a number of very profitable data-driven media companies, who have simply applied this principle earlier than others. They enjoy healthy growth rates and sometimes impressive profits. One example: Bloomberg. The company operates about 300,000 terminals and delivers financial data to its users. If you are in the money business this is a power tool.”³⁶

While the best examples of data supporting media business are still from Northern countries,³⁷ Craig Hammer of the World Bank Institute suggests that there are opportunities already for media houses in other countries to start by capitalizing on their own data, for example, scraping the data archived in Kenya’s *The Nation*. “They are sitting on what could be actionable intelligence” that no one has yet tapped into, he explained.³⁸

The Reuters Institute for the Study of Journalism at Oxford’s 2013 Digital News Report—which looked at the U.K., the United States, Germany, France, Italy, Spain, Brazil, Japan, and Denmark—noted, “There are marked differences with countries where populations were late adopters of online services or in which the legacy of free online news provision is less pronounced. The highest rates of paying for digital news are thus seen in Brazil, Italy, Spain, Japan, and France.” People are accessing news more frequently through new devices, and the study found increasing revenue from digital apps, not just pay walls. Significantly, the youngest people in each country were the most willing to pay for news, which seems to signal a promising future for paid content. A significant percent of those who are not paying now expect to pay in the future. In Brazil— with the highest number of such willing subscribers in the study—almost 60 percent of all news users expect to pay for news in the future.³⁹ In some countries at least, the future is showing signs of revenue from digital news, which should help to make the case that data journalism can contribute to the financial strength of the media.

Building skills: beyond training

Data journalism relies upon skills that are not usually found in most newsrooms, and most universities and training programs do not offer courses on the topic. Building journalists’ data skills is clearly challenging. Even the most successful data journalists are continually improving their skills, because digital tools are always evolving. What would be the best approach to building and increasing data skills? The media development literature is overloaded with lessons cautioning that short-term, parachute training engagements have limited to no impact. Some form of long-term educational engagement designed specifically for adult learners will be necessary for data journalism.

Boot camps have been a favored approach to quick, intensive skill building in data journalism, but boot camps can only be effective if they are not done in a vacuum. The boot camps also have their skeptics. Are all boot camps really that intensive? Is even a week of intense focus really enough? More importantly, what happens after journalists get back into the newsroom? The realities of everyday journalism often clash with and override what can be learned in an intensive boot camp; this too has been a commonly articulated lesson from media development. Without building institutional support in the newsroom, data journalism cannot grow. Current thinking is that trying to offer different approaches, with a more long-term strategy, that tie training activities to education, professional development, and other forms of civic engagement would yield better results. (See Appendix A, Case Studies.)

Open source education and tools

Another avenue open to media developers that has been helping to spread data journalism resources and knowledge, and to some extent build networks, is Massive Open Online Courses, or MOOCs. These online courses, usually taught through a university and offered free of charge to anyone with an Internet connection, have grown significantly in the last two years. MOOC instructors use text, video, visuals, topic-driven conversation threads, and often require electronically administered quizzes and assignment submission. The Knight Center for Journalism in the Americas at the University of Texas has offered several classes just in the last year, in English and in Spanish, on data journalism and data visualization. Argentina's Crucianelli taught an introduction to data journalism in the spring of 2013, which attracted almost 4,000 students from 62 countries. According to Knight Center Director Rosental Alves, the suite of courses offered were the first global MOOCs on journalism.⁴⁰

The European Journalism Centre (EJC) offered a free online data journalism course in May 2014 entitled "Doing Journalism with Data: First Steps, Skills and Tools," which it billed as "the largest massive open online course (MOOC) on the topic of journalism to date." The course was part of the EJC's "Data Driven Journalism initiative, which aims to enable more journalists, editors, news developers and designers to make better use of data and incorporate it further into their work."⁴¹ During the same time period, Google offered a free course, "Making Sense of Data," which "is intended for anyone who wants to learn more about how to structure, visualize, and manipulate data," including journalists. More free online courses are certain to follow.

The degree to which MOOCs will be successful in teaching students a complex topic like data journalism remains to be seen. While the MOOC format is not completely self-taught, it still requires a high degree of investment of time and focus, and students can feel isolated, despite some conversational interaction among them and the ability to review other's assignments. MOOCs do not offer a strong support community, feedback loop, or individual support. In addition to MOOCs, other institutions and universities such as Poynter's News University offer shorter, free webinar workshops on topics related to data journalism. The Open Knowledge Foundation's School of Data offers a free online toolkit for individuals to join or lead a "Data Expedition."⁴²

Other free online resources include the *Data Journalism Handbook*, a collaborative, open source, online text.⁴³ The 71 contributors overwhelmingly hail from the developed world, but include a handful of people from the Buenos Aires chapter of Hacks/Hackers, and a few others from the Global South. An Iberoamerican data journalism handbook is also under development. Rather than a translation of the original, this volunteer project is meant to address the Latin American context more specifically.⁴⁴

In addition to online classes and texts, many of the online tools that journalists use for work with data are free, such as the Google suite of apps, Tableau Public, source code available on Github, and numerous others. Many of these tools were developed under an open source model, wherein developers self-organize, collaborate, and barter, rather than work under a commercial model. The producers of these tools maintain an ethos of volunteerism and sharing of skills and resources. So, too, many of the best data journalists share content, rather than jealously guarding a scoop.

Free online learning resources are unlikely by themselves to be responsible for spawning a whole generation of well-skilled data journalists. However, together, the range of networks, tools, and texts offer opportunities for journalists with Internet access to build and grow their knowledge about data. Importantly, they also offer a community of potential colleagues with whom to collaborate. These various resources, collaborations, and alternative ways of learning hint at a shift in power; journalists may have the opportunity to become data literate and part of a broader community of data journalists, even if their immediate environment does not nurture it. All of these additional tools and guides offer increased resources for media developers to put in place further measures of support for trainees and to help sustain and build knowledge and skills long after a particular intervention. (For more, see Appendix B:Online at (http://cima.ned.org/sites/default/files/CIMA-Data%20Journalism_Resources.pdf.)

Networks and collaboration

From a media development perspective, the transformation of journalistic work to include a reliance on open source resources and loose, flexible networks of collaborators has valuable potential. Better access to resources, partners, and mentors may increase impact of training and other interventions.

Hacks/Hackers is a prime example of a volunteer network of collaborators on data journalism. An informal organization with chapters around the world, Hacks/Hackers is an open-source organizational idea for people interested in data journalism and visualization to gather and work together. Miguel Paz, who co-founded the Hacks/Hackers chapter in Santiago, Chile, said:

The cool thing about it, it's an organization that's not an organization. It's a set of rules, and if you apply and agree on the rules, you can do it. It's not owned by any one media group. Journalists from one news group usually won't go to things sponsored by another one. It's a temporary free zone to meet other journalists, people that come from a start-up background, data scientists, sociologists interested in news ... When they go to hackathons, they realize that journalists, programmers, designers are brainstorming on how to work together. It lowers the barriers of distrust.⁴⁵

A posting on the Hacks/Hackers website on September 5, 2013, acknowledged the largest gathering ever of a Hacks/Hackers, in Buenos Aires, with close to 1,000 participants.⁴⁶ The website currently lists about 30 countries with chapters (some countries have multiple chapters across different cities). Cui Anyong said that he had recently applied for a chapter in Beijing, signaling a spark of hackathon culture beginning to grow in China.

Beyond the Hacks/Hackers organization, many other hackathons that either focus on journalism or have journalists participating are also creating spaces to jumpstart data-driven projects where there may be insufficient capacity, time, or interest within the media houses themselves. Even when there is a focus on data journalism within a media outlet, hackathons provide the opportunity for new kinds of collaboration and knowledge exchange—all voluntary—that invite a different kind of cross-fertilization of ideas than what happens within just one organization. In Hong Kong, an inaugural “catalyst night” was meant not just to be a one-off intense hackathon, but the beginning of a long-term network of people that would work on more sustainable projects.⁴⁷ Open News has organized between 50-60 journalism-themed hackathons around the world, “in every

continent but Antarctica,” according to Dan Sinker of Knight-Mozilla OpenNews. These events help identify people who are interested in the intersection of programming and journalism, but “I believe hackdays are a sleight of hand,” Sinker said. “They don’t produce useful code . . . What hackdays are good for is identifying people and building community.”⁴⁸

There are of course many challenges to working in networks, even in countries where journalists have access to the Internet. But even in closed or authoritarian regimes, journalists can use networks to follow corruption and crime across borders. Paul Radu, who directs the Organized Crime and Corruption Reporting Project (OCCRP), based in Romania, argues that “even in the most austere environments,” access to information located outside of the country, together with new technology, and networks of collaborators make cross border collaborative investigations possible.⁴⁹

The increase in collaborative, multi-disciplinary, sometimes transnational collaboration visible across data journalism heralds a transformation of stovepiped labor institutions into flexible, changing, networked spaces. In *Big Data*, Mayer-Schonberger and Cukier write, “In the spirit of Google or Facebook, the new thinking is that people are the sum of their social relationships, online interactions and connections with content. In order to fully investigate an individual, analysts need to look at the widest possible penumbra of data that surrounds the person—not just whom they know, but whom those people know too, and so on.”⁵⁰

Collaboration in data journalism by nature includes non-journalists; it may begin to include citizen journalists as well. Giving a broad vision for the future of U.S. journalism, Tom Rosenstiel, executive director of the American Press Institute, argues, “Journalism’s future must be a collaboration in which citizens, technology and professional journalists work together to create a public intelligence that is deeper and wider than any of these could produce alone.”⁵¹ Journalism, particularly data journalism, is likely to become increasingly collaborative and networked across the globe.

Experimentation and innovation

The rapid change of information systems in recent years, together with the destabilization of journalism, present good opportunities for experimentation in how best to develop data journalism. Hacker culture by nature is experimental; those that came into data journalism from the ICT side seem particularly comfortable with an entrepreneurial approach. Before entering into data journalism, Darcy Christ of Hong Kong University worked on

the content management system of an online magazine, where he watched the changes in the journalism field close up: “This industry was seemingly stable but really challenged. I got excited; it seemed like no one had a plan. Everything was changing—the stories we tell, how we were telling the stories, who was telling the stories—and data journalism is all about that. People pushing things because they think something interesting will come along.”⁵² The insecurity about rapid change with an unclear future was exciting to the hackers.

Contemporaneous with the rise of data journalism, and sometimes used in concert with it, are other new journalistic forms evolving out of experimentation and innovation—digital platforms for reporting, journalistic apps, and citizen journalism. All of these new developments are still finding their footing and place in the world, but they can be seen as “different pieces of the same puzzle,” according to Miguel Paz.⁵³ Figuring out how to successfully put the different pieces together is still in the experimental stage, an appropriate moment for media developers to try out small-scale initiatives. Several interviewees pointed to Knight International journalism fellow Justin Arenstein as an example of a journalist who is attempting to construct a broad vision for journalism by synthesizing all of these different pieces. Arenstein, according to the International Center for Journalists’ website, “is serving as an ideas catalyst and ‘match-maker’ by leveraging his deep links within the African media industry and his entrepreneurial networks to generate disruptive ideas.”⁵⁴ Working with the Africa Media Initiative, Arenstein’s projects include the African News Innovation Challenge; a Prototype Fund; Code4Africa; Hacks/Hackers across Africa; and others.

Conclusion

The rapid expansion of digital data provides an extraordinary opportunity to improve evidence-based governance and to increase citizen participation in decision making. Yet without media that can make sense of this growing mountain of data, the promise of the transparency revolution is likely to be lost.

Acknowledgement of the increasingly central role of data in decision making at all levels of society is increasingly visible. The High-Level Panel on the Post-2015 Development Agenda called for “A New Data Revolution” that would help track progress toward development goals and ensure the inclusion of all people in human development.⁵⁵ But for data to truly forge inclusiveness in development, multilateral organizations, governments, NGOs, companies, and citizens will increasingly need “infomediaries” to chase down and make sense of the most relevant data of interest to people. This new data can be an important building block for creating sustainable media institutions, stimulating wider demand for fact-based policy and decision making, and measuring progress. Today, more transparency in budgets, spending data, or service provision statistics can likewise be a critical raw material for enterprising media.

International development also needs stronger data journalism to investigate its successes and failures and to improve the targeting and effectiveness of aid interventions. For example, an initiative called AidData provides Web-based tools and resources for analyzing aid flows to determine the impact of donor aid, identify gaps, and plan future investments.⁵⁶ Journalists are needed to analyze and interpret for citizens whether aid has been effective and how investments relate to their own communities. The transparency and accountability promotion community could increase the effectiveness of its work by supporting the growth and uptake of this type of data journalism.

But all of these promises will not be realized just by training journalists and providing them with the latest digital tools. The specter of all of the other well-known challenges to practicing journalism—censorship, attacks on journalists, criminal libel laws, and collapsing business models—is a reminder that in the absence of a stable, enabling, and supporting environment, data journalism is likely to remain an unfulfilled promise.⁵⁷ The international development community should work more closely with media developers to ensure that the critical role of media is well understood and factored into overall development planning. More openness and transparency can be a major contribution to the success of development. A strong, independent, and capable media is central to this vision.

Appendix A: Case Studies

How are media developers meeting these challenges? Case studies

In recent years, the media development community has begun to integrate data journalism into implementation initiatives. The 2012 Global Forum for Media Development conference recognized this, and had an observable emphasis on data journalism, “to look at if and how it changes news reporting, how journalists and other information providers (NGOs, campaigners etc.) can make use of it and how it fits into the media development strategy.”¹

We encountered several different models for building data journalism capacity, and six are featured here: IRE’s NICAR conferences and trainings; the Journalism and Media Studies Centre at Hong Kong University; Knight International Journalism Fellowships; Internews in Kenya’s Digital Media Center; World Bank Institute’s Global Media Development Program; and the Knight-Mozilla Fellowships. This diversity of approach stems both from the vastly different contexts in which data journalism arises, and the creative and experimental spirit that is characteristic of its practitioners.

1. Bettina Peters, email with author, March 9, 2013

IRE - NICAR: Pioneer and Connector

Brief description

As an educational institute and a network, no other organization compares to IRE's NICAR, which runs trainings, holds international conferences, maintains a database, and produces and offers resources like books, webinars, and tipsheets.

"The National Institute for Computer-Assisted Reporting is a program of Investigative Reporters and Editors, Inc. and the Missouri School of Journalism. Since 1989, NICAR has trained thousands of journalists in the practical skills of getting and analyzing electronic information and also maintains a library of databases containing government data on a wide array of subjects, available for purchase by IRE members," according to its website. The organization has held many international conferences and trainings over two decades, and trained many of the journalists who went on to become data pioneers throughout the globe. In early 2013, NICAR held its largest conference ever (more than 600 attendees). The spike in attendance in recent years has been attributed to the entrance of people with a computer background. <http://www.nicar.org>

What makes it unique

NICAR has been present from the early days of personal computers in the newsroom up through the boom in data journalism. Parent organization IRE is located at the Missouri School of Journalism, which provides cross-fertilization and support.

Examples

A three-day boot camp on "Building a web scraper" took place in mid-October 2013 at the University of Missouri campus.

The most recent annual CAR conference took place over four days in February-March 2014 and featured numerous hands-on workshops.

Potential impact

NICAR has a substantial reach and is well known. It is said to have originated the "boot camp" approach; hands-on training over a short period of time (usually a few days). Conferences keep people connected after the boot camps. The organization has been a strong force for community-building as well.

Challenges

NICAR was the institutional early pioneer in data journalism, but it is behind the curve in the open data movement, which is an important support to data journalism. NICAR's data is only available to IRE members, although some of it is available for purchase. Its adherence to the moniker "Computer-Assisted Reporting" gives it the appearance of being a bit of a throwback. However, these challenges are minor; the organization is well respected and has trained many of the field's greatest achievers.



International Center For Journalists (ICFJ) Knight Fellows: Networked Changers

Brief description

In 2007, the Knight fellowships were redesigned from a kind of Peace Corps for journalists towards longer-term project-based fellowships, lasting at least a year. ICFJ began to increasingly include journalists from outside the United States, and focused on countries where it believed that fellows could have the greatest impact. The fellows' work focuses in one or more of four areas: data, mobile, social networks, and multimedia storytelling. Several of the fellowships, funded by the Bill and Melinda Gates Foundation, focus on Africa. In recent years, the program has focused the rest of the fellowships on Latin America and on journalists working with data. Although the fellows have separate projects and work individually in different countries, there is a high degree of collaboration among them, both under the fellowships and beyond.

What makes it unique

The strength of the network and the degree of the fellows' collaboration on an array of different types of projects is unique.

The very deliberate re-design of the program to increase impact, followed by a public evaluation, is notable.

Examples

Miguel Paz introduced lo-fi technology to Chile that is being used for rural community radio in Africa. Exposure to fellows based in Africa gave him the idea to use a technology that works in settings where infrastructure and capacity is limited.

Current fellow Paz and former fellow Sandra Crucianelli are collaborating on the Iberoamerican Data Journalism Handbook.

Potential impact

The program began to focus more on being able to measure impact on media organizations, and in turn, the organizations' impact on society. A 2011 report documented tangible impacts achieved by the new approach to the fellowship. This included an emphasis on recruiting entrepreneurial fellows with management and language skills, in addition to a journalistic background.

Challenges

Daily demands, limited resources, and the ever-increasing rapidity of journalistic practice are barriers to fellows' impact. Creating a culture of news innovation is a significant challenge.

Internews in Kenya: Deep Focus in One Country

Brief description

Internews realized after the well-publicized 2011 launch of the Kenya Open Data Initiative, that no stories were being produced with the newly available data. In Kenya, data went from being completely closed to completely open, so there is a gap in journalists' understanding of how dynamics work in even getting experts to comment on data. Journalists prior to this had no access to data, so they struggled quite a bit when there was sudden access. Internews created a focus on data journalism to bridge the gap between open data and a lack of numeracy in Kenyan journalists. Offered a grant by the Population Reference Bureau, Internews began to teach journalists to attach numbers to trends and use simple language. As of 2013, a component of each of their grants focuses on data journalism.

What makes it unique

Internews set up a Digital Media Center, a media resource center originally created out of health journalism funds. It became a space to train and mentor journalists on a range of digital journalism skills, including working with data. The Digital Media Center has an ongoing presence, full-time staff, and is meant to be the opposite of parachute training. It is a resource for both individual journalists and media houses.

Example

The Data Dredger, a portal to boost Kenyan journalists' use of data in their reporting, "a space where journalists can mine data that is packaged in an user friendly manner—ready to be used in storytelling. Storytellers can also download compelling visualizations for their stories."

<http://www.internewskenya.org/dataportal/content?page=about>

Potential impact

Dozens of stories produced that brought new issues to light, such as the imbalance in Kenyan exports (which have remained flat for 30 years) versus imports, which have tripled. Journalists trained by Internews have produced many stories for which they found data and produced analysis that the experts (economists, health researchers, etc.) did not have.

Challenges

There is a high level of innumeracy among Kenyan journalists. Even seasoned journalists do not understand statistics or the scientific method. There is no culture of understanding how to talk about numbers, how to interview a scientist, or how to determine if data makes sense. Also, much of the opened data was not in good shape to work with, and the Open Data Initiative is voluntary, so only limited data is available.



World Bank Institute (WBI): Connecting the Dots

Brief description

Throughout the World Bank, the only program completely dedicated to media development is WBI's Global Media Development Program. Staffed by one person, Craig Hammer, the program is entirely focused on data journalism. The Global Media Program sits at the intersection of WBI's emphasis on open government, open knowledge, open data, and citizen participation in governance. WBI neither conducts training nor funds it but has a role as facilitator, working with different stakeholder groups to come together around issues and create collective goals. Considering the importance of networks, collaboration, and multi-disciplinary teams in data-driven journalism, this role as a connector is important.

WBI's approach emphasizes the boot camp model of building skills, using hands-on training in data skills over a short period of time (up to a week). It is also helping to feed training materials used in the boot camps to journalism schools. It also encourages media organizations to think more about their own data to better target audiences. WBI works with Hacks/Hackers, and also works with CSOs to create demand for data, engage with data, define public priorities, and demand better services.

What makes it unique

The status and reach across countries and organizations conferred by the World Bank puts the program in a unique situation, despite its small size. Its work in media has been demand driven, fielding requests from World Bank country teams, as well as journalists and media houses. It works with most of the other organizations featured in the case studies.

Example

South-South knowledge exchange—sharing information in training with a similar cohort in another country—is highly emphasized. Data journalists and scientists from Kenya went to Moldova in 2012 to lead a boot camp, teaching Moldovan journalists specific skills and imparting more broadly what data journalism means and the impact it can have.

Potential impact

WBI is a relationship broker and knowledge broker in a field that amplifies the strength of networks. It seems to have taken some key lessons from media development to heart; "Continuing engagement is the most important part of the process; a community of practice can last. Training is momentary," according to Craig Hammer.

Challenges

The World Bank has had a changing relationship to media development. It is not clear that the bank fully recognizes the role that media has to play in the development process. Impact could be much greater with more support and recognition from the Bank.

Knight-Mozilla Fellows: Starting from a Tech Perspective

Brief description

The Knight-Mozilla fellows, like the ICFJ fellows, include a broader range than just data journalism, but make a significant contribution to it. The first fellowships began in 2011 to place “developers, technologists, civic hackers, and data crunchers” in newsrooms as change agents, for open source development. When Dan Sinker joined the organization, he broadened the scope and re-launched in 2012 as OpenNews. The new organization selects and places fellows, sponsors hack days, runs a website called Source, documenting work that’s happening in the journalism community, (including a learning section), and sponsors “code sprints” to expand the code available to people doing back end tools for journalism.

Fellows, who are already quite accomplished in their fields, spend 10 months “hacking the news.” They build resources within newsrooms and strengthen already strong capacities of multi-disciplinary teams within newsrooms.

What makes it unique

The range of activities goes beyond the scope of data journalism, but illustrates very well the broader set of activities and relationships that nurture it. OpenNews also supports the enabling environment in which data journalism can flourish.

Example

Open “community calls” allow the public (via conference call) to hear about the projects and interact with the fellows and become part of the journalism code community.

Potential impact

All projects the fellows work on are open source, and their mandate includes sharing the code they write and blogging about their experiences. Fellows are meant to strengthen relationships between journalists and developers and contribute to resources and transparency on the internet.

Challenges

Fellows are embedded at a range of news outlets, most of which have been from the Global North, and all of which appear to already have strong capacity in data-driven journalism and digital journalism (e.g. The New York Times, The Guardian, ProPublica, La Nacion (Argentina), etc.). This likely increases what the fellows are able to produce, and is a strategy likely to give the fellows a great experience, but does not strengthen struggling news outlets.



**Hong Kong University:
Data Journalism embedded in, and practiced at,
the Journalism and Media Studies Centre**

Brief description

Ying Chan, founder and director of the journalism school at HKU, introduced data journalism into the curriculum in 2011. The faculty also practice data journalism. The approach to teaching data journalism? “The core values are the same – we teach them to look for facts, information, stories; to look for stories that serve the public interest, to try to deliver them in ways that are interesting,” according to Ying Chan.

The JMSC is experimenting with models to figure out the best way to train data journalists; currently the model is to recruit computer scientists to learn about journalism. Google gave three scholarships to students with a background in computer science to study journalism at HKU. Faculty also tries to encourage students to become skilled at data journalism, arguing that it will make them more employable.

Professors and students are engaged with data outside of the classroom, contributing to the Data Journalism Lab, participating in hackathons, advocating for the Hong Kong government to open its data, among other activities.

What makes it unique

Many interviewees complained that in their countries, the journalism schools don’t offer any data journalism classes. Multi-disciplinary faculty members teach the range of skills needed. It attracts well-known guest instructors, such as Irene Jay Liu and Jonathan Stray.

Example

The Data Journalism Lab is in a pilot phase; the team includes a mix of faculty and students. About the Lab: “This platform builds a data centre and a set of interactive websites to empower citizens in Hong Kong to participate in various issues regarding governance, politics and new developments. Citizens in Hong Kong usually have limited access to fragmented and technical documents and reports. Here data from multiple government departments will be scraped, cleaned and transformed into visual information.” <http://datalab.jmsc.hku.hk/>

Potential Impact

Many students will work in Hong Kong and China after graduation; there is potential for impact on China from both locations. The data perspective on China is potentially an alternative to the widespread government and corporate perspectives. The faculty is advocating for a Freedom of Information Act in Hong Kong, for its own good, and also driven by the idea that getting a Hong Kong FOIA could have a positive impact on China.

Challenges

In Hong Kong, people spend a lot of time online, but not there is no strong interest in online social networks. Darcy Christ observes that Hong Kong, with 7 million people, still operates in many ways by word of mouth (amplified by high mobile phone use). Hong Kong is perhaps behind expectations in terms of innovation.

Endnotes

1. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/207772/Open_Data_Charter.pdf, Accessed July 21, 2013.
2. See <http://globalopendatainitiative.org/>.
3. <http://www.opengovpartnership.org/how-it-works/requirements>, Accessed, July 21, 2013.
4. Steve Doig, interview with the author.
5. For a discussion of scoring methodology, see “Open data Barometer: 2013 Global Report,” a joint project of the World Wide Web Foundation and the Open Data Institute: <http://www.opendataresearch.org/dl/odb2013/Open-Data-Barometer-2013-Global-Report.pdf>, p. 21. Chart based on bar graph in same report.
6. Brant Houston, interview with the author.
7. World Economic Forum, “Big Data, Big Impact: New Possibilities for International Development,” http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf, p.3, Accessed July 21, 2013.
8. Steve Lohr, “The Age of Big Data,” *New York Times*, February 11, 2012, <http://www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html?pagewanted=all>, Accessed July 21, 2013.
9. Raw data at <http://www.itu.int/en/ITU-D/Statistics/Pages/stat/default.aspx>, Accessed July 21, 2013. Note that for the ITU, individual internet access is defined by use “in the last 12 months” and so paints a very impressionistic picture of actual use. See http://www.itu.int/en/ITU-D/Statistics/Documents/publications/manual2009/CoreICTIndicators_e_rev2.pdf for indicator definition. Nonetheless, the upward trend is observable if perhaps overstated in magnitude.
10. See for example, <http://www.guardian.co.uk/technology/2013/jun/15/database-app-icij-tax-havens>.
11. <http://www.icij.org/offshore/secret-files-expose-offshores-global-impact>, Accessed September 7, 2013.
12. <http://www.icij.org/blog/2013/06/how-we-built-offshore-leaks-database>, Accessed September 7, 2013.
13. Irene Jay Liu, “Welcome to Connected China,” February 28, 2013, <http://blogs.reuters.com/connected-china/2013/02/28/welcome/>, Accessed September 7, 2013.
14. Miguel Paz, interview with the author.
15. Jose Toledo, interview with the author.
16. See <http://www.journalism.co.za/images/videos/DiariosSecretos.mov>; <http://www.rpctv.com.br/diariossecretos-english/>.
17. Committee to Protect Journalists, Journalists Killed by Year, <http://www.cpj.org/killed/2013/>, Accessed February 2, 2014.
18. Doig, Interview with the author.
19. Eva Constantaras, “Open Sesame Does Not Work for Open data,” <https://innovation.internews.org/blogs/>

[open-sesame-does-not-work-open-data](#), Accessed 2/27/13.

20. Alex Howard, "The growing importance of data journalism," December 21, 2010, <http://strata.oreilly.com/2010/12/data-journalism.html>, Accessed August 24, 2013.
21. "Open data initiative being frustrated, says Matiang'I" Daily Nation, May 29, 2013 by Nation Correspondent <http://www.nation.co.ke/business/news/Open-Data-Initiative-being-frustrated-says-Matiangi/-/1006/1866216/-/menwxz/-/index.html>.
22. Eva Constantaras, Interview with the author.
23. Mark Horvit, Interview with the author.
24. Elisa Tinsley, Interview with the author.
25. Sandra Crucianelli, Interview with the author.
26. Segnini, Interview with the author.
27. For example, the 2012 Code4Africa pilot program, see www.code4kenya.org. Eva Constantaras' research found that the program was attempting to create news apps in an overly limited time period, before establishing basic data literacy and good, institutionalized working relationships.
28. Ida Jooste, Interview with the author.
29. See <http://www.latimes.com/news/local/la-me-geocoding-errors5-2009apr05,0,5966285.story> for more information.
30. Brian Caffo, professor of biostatistics at the Johns Hopkins Bloomberg School of Public Health, quoted in Carl Bialik, "Upbeat stats on Statistics," <http://blogs.wsj.com/numbersguy/the-upbeat-stats-on-statistics-1216/?mod=WSJBlog>, Accessed August 25, 2013.
31. Dan Sinker, Interview with the author.
32. Michael Zanchelli with Sandra Crucianelli, "Integrating data journalism into newsrooms," ICFJ, n.d. http://www.icfj.org/sites/default/files/integrating%20data%20journalism-english_0.pdf; the authors found a similar phenomenon at the level of the newsroom: "We found that each data-journalism team could point to a story that was a sort of "game-changer" for the organization. These stories were often extremely personal, connecting data with the lives of their audience," p.6.
33. Mirko Lorenz, "Business models for Data Journalism," in the Data Journalism Handbook, http://datajournalismhandbook.org/1.0/en/in_the_newsroom_10.html, Accessed September 7, 2013.
34. Mallery Jean Tenore, Texas Tribune databases drive majority of site's traffic, help citizens make sense of government data, Mar. 2, 2011, <http://www.poynter.org/latest-news/top-stories/121281/texas-tribune-databases-drive-majority-of-sites-traffic-help-citizens-make-sense-of-government-data/>, Accessed September 7, 2013.
35. David Donald, Center for Public Integrity, interview with the author.
36. Lorenz, Ibid.

37. Though as with the data journalism field overall, we are still at the beginning even in the Northern Countries. Nicolas Kayser-Bril offers several interesting ideas on “How to Make Money with Data Journalism in the US and Europe” that may eventually apply to other countries as well. See [http://datadrivenjournalism.net/news_and_analysis/How to Make Money with Data Journalism in the US and Europe](http://datadrivenjournalism.net/news_and_analysis/How_to_Make_Money_with_Data_Journalism_in_the_US_and_Europe).
38. Craig Hammer, Interview with the author.
39. From Reuters Institute for the Study of Journalism, Nic Newman, report editor. <http://www.digitalnewsreport.org/>, Accessed July 2, 2013.
40. See <https://knightcenter.utexas.edu/00-13876-knight-centers-mooc-introduction-data-journalism-starts-3000-students-and-high-participatio>.
41. The European Journalism Centre, press release, March 5, 2014.
42. See <http://schoolofdata.org/data-expeditions/>.
43. See <http://datajournalismhandbook.org/>.
44. Paz, interview with the author. Crucianelli, interview with the author. Paz is spearheading the effort and Crucianelli will be the editor.
45. Paz, interview with the author.
46. See <http://hackshackers.com/>, Accessed, September 7, 2013. Also see <http://blogs.dw-akademie.de/africa/?p=9015> Kate Hairsine interview of Justin Arenstein, for information on Hacks/Hackers in Africa.
47. See <http://baoman.wordpress.com/2013/05/14/open-data-hong-kong-catalyst-night/>
48. Sinker, interview with the author.
49. Paul Radu, “Following the Money: cross-border collaboration,” in the Data Journalism Handbook, http://datajournalismhandbook.org/1.0/en/in_the_newsroom_7.html.
50. Quoted in Mayer-Schonberger and Cukier, *ibid*.
51. Tom Rosenstiel, “The danger of journalism that moves too quickly beyond fact,” June 17, 2013 <http://www.poynter.org/latest-news/the-next-journalism/216034/the-danger-of-journalism-that-moves-too-quickly-beyond-fact/>, Accessed July 1, 2013.
52. Darcy Christ, interview with the author.
53. Miguel Paz, interview with the author.
54. Knight Fellowship profile: <http://www.icfj.org/our-work/sub-saharan-africa-spurring-innovation-and-experimentation-newsrooms>, Accessed July 1.
55. A New Global Partnership: Eradicate Poverty and Transform Economies Through Sustainable Development. The Report of the High-Level Panel of Eminent Persons on the Post-2015 Development Agenda. United Nations, 2013, p. 23. http://www.un.org/sg/management/pdf/HLP_P2015_Report.pdf.
56. See <http://open.aiddata.org/content/index>.



57. See Mark Nelson with Tara Susman-Peña, “Rethinking Media Development: A Report on the Media Map Project,” for an extensive evidence-based discussion of this issue. <http://www.mediamapresource.org/research-and-resources/>.

Advisory Council for the Center for International Media Assistance

Esther Dyson

Craig LaMay

Stephen Fuzesi, Jr.

Caroline Little

William A. Galston

William Orme

Suzanne Garment

Dale Peskin

Mark Helmke

Adam Clayton Powell III

Ellen Hume

Monroe E. Price

Jerry Hyman

Adam Schiff

Alex S. Jones

Kurt Wimmer

Shanthi Kalathil

Richard Winfield

Susan King

Center for International Media Assistance

National Endowment for Democracy

1025 F Street, N.W., Suite 800

Washington, DC 20004

Phone: (202) 378-9700

Fax: (202) 378-9407

Email: CIMA@ned.org

URL: <http://cima.ned.org>



**National Endowment
for Democracy**

Supporting freedom around the world