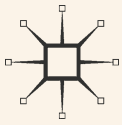
DIGITAL INVESTIGATIVE JOURNALISM

Data, Visual Analytics and Innovative Methodologies in International Reporting

EDITED BY
OLIVER HAHN AND FLORIAN STALPH



Digital Investigative Journalism

Oliver Hahn · Florian Stalph
Editors

Digital Investigative Journalism

Data, Visual Analytics and Innovative Methodologies in International Reporting

palgrave
macmillan

Editors
Oliver Hahn
Centre for Media and Communication
University of Passau
Passau, Germany

Florian Stalph
Centre for Media and Communication
University of Passau
Passau, Germany

ISBN 978-3-319-97282-4 ISBN 978-3-319-97283-1 (eBook)
https://doi.org/10.1007/978-3-319-97283-1

Library of Congress Control Number: 2018950725

© The Editor(s) (if applicable) and The Author(s) 2018
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Cover credit: wektorygrafika
Cover design by Tom Howey

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

1 Digital Investigative Journalism 1
Oliver Hahn and Florian Stalph

Part I Data-Driven Investigation and Computational Journalism

2 The Rise of Investigative Data Journalism 9
Andrew W. Lehren

3 Zeroes and Ones: Investigating with Data 19
Paul Bradshaw

4 Measuring the Unmeasured with Data 31
Nicolas Kayser-Bril

5 Understanding and Contrasting the Data and Exploratory News Outlets 43
Felipe De La Hoz

6 Uncovering International Stories with Data and Collaboration 55
Emilia Díaz-Struck and Mar Cabra

7 Data-Driven Human Rights Investigations 67
Alice Kohli

8 Following the Money Trail: Investigative Data Journalism 79
Miranda Patrucic

Part II Immersive Storytelling and Visual Investigation

9 Drone Journalism: Storytelling from a New Perspective 91
Ben Kreimer

10 Immersive Journalism: How Virtual Reality Impacts Investigative Storytelling 103
Marcus Bösch, Stephan Gensch and Linda Rath-Wiggins

11 Visual Storytelling: Show, Not Tell? Show AND Tell 113
Jens Radü

12 Interactivity to the Rescue 123
Johanna Fulda

13 Visual Analysis: Verification via Geolocation and Photographs 137
Eliot Higgins

14 eyeWitness to Atrocities: Verifying Images with Metadata 143
Eleanor Farrow

15 A Matter of Perspective: Truth, Evidence and the Role of Photography as an Investigative Tool 157
CJ Clarke

16 The Context Verification of Digital Photographs 171
Alexander Godulla

17 Photo Manipulation: Software to Unmask Tampering 179
Alessandro Tanasi

Part III Authenticity, Identity and Transparency

18 Fact-Checking as Defence Against Propaganda in the Digital Age 193
Anna Sarmina

19 Crowdsourced and Patriotic Digital Forensics in the Ukrainian Conflict 203
Aric Toler

20 In-Depth Crisis Reporting 217
Natalia Antelava

21 eJihad: Behind the Use of Social Media by ISIS 229
Asiem El Difraoui with Oliver Hahn

22 Truth Corrupted: The Role of Fact-Based Journalism in a Post-Truth Society 237
Florian Stalph

23 The Future of Investigative Journalism in an Era of Surveillance and Digital Privacy Erosion 249
Julie Posetti

Index 263

Notes on Contributors

Natalia Antelava is a co-founder and editor-in-chief of Coda Story.

Marcus Bösch is a journalist, lecturer for mobile reporting, newsgames and virtual reality and a co-founder of Vragments.

Paul Bradshaw runs the data journalism and multiplatform and mobile journalism master programmes at the School of Media, Birmingham City University, UK. He also works as a data journalist with the BBC data unit.

Mar Cabra is a board member of the Global Editors Network and member of the International Consortium of Investigative Journalists (ICIJ), where she was head of the data and technology team until September 2017. She is one of the pioneers of data journalism in Spain.

CJ Clarke is an award-winning film director and photographer. He is the author of "Magic Party Place", a ten-year project mapping the rise of the right and the roots of Brexit. The book was shortlisted for the Aperture Paris Photo First Book Award.

Felipe De La Hoz is a freelance politics and immigration journalist based in New York City and writing regularly for the Village Voice and Gotham Gazette, among others.

Emilia Díaz-Struck is a research editor at the International Consortium of Investigative Journalists (ICIJ).

Asiem El Difraoui, Ph.D. is a political scientist, economist, and documentary filmmaker. He is also a researcher at the Institut d'études politiques de Paris (Sciences Po), France, and a senior fellow at the Institut für Medien- und Kommunikationspolitik in Cologne, Germany.

Eleanor Farrow is the Project Coordinator for the eyeWitness to Atrocities project under the auspices of the International Bar Association in London.

Johanna Fulda is front-end developer and data visualisation designer at Cumul8. She was a member of the InfoVis Group at the University of British Columbia, Canada.

Stephan Gensch is a journalism technologist and head of product development and a co-founder of Vragments.

Prof. Dr. Alexander Godulla is Professor of Empirical Communication and Media Research at the University of Leipzig, Germany.

Prof. Dr. Oliver Hahn is Professor of Journalism at the University of Passau, Germany.

Eliot Higgins is the founder of Bellingcat. He is a visiting research associate at the King's College, London, UK, and visiting research fellow at the Human Rights Center at the University of California, Berkeley.

Nicolas Kayser-Bril is a data journalist, and was a co-founder and manager of Journalism++ from 2011 to 2017, and former head of data journalism at Owni.

Alice Kohli is a researcher for Public Eye in Switzerland and a former data journalist at the *Neue Zürcher Zeitung*.

Ben Kreimer is an independent journalism technologist and advisor for the Drone Journalism Lab. He was BuzzFeed's first Open Lab fellow, and has worked with many academic institutions and organisations, including Columbia's Tow Center for Digital Journalism, USC Annenberg, CCTV Africa, *The Times of India*, and the *USA TODAY* Network.

Andrew W. Lehren is a senior investigations editor for NBC News, specialising in data journalism. He was a long-time reporter for *The New York Times*. He is the director of investigative reporting at the City University of New York's Graduate School of Journalism.

Miranda Patrucic is an award-winning investigative reporter and regional editor for the Organized Crime and Corruption Reporting Project (OCCRP).

Julie Posetti is a multi-award-winning Australian journalist and journalism academic who is a senior research fellow with the Reuters Institute for the Study of Journalism at the University of Oxford, UK. She is the author of UNESCO's *Protecting Journalism Sources in the Digital Age*.

Jens Radü is a journalist and Head of Multimedia at *Der Spiegel*.

Linda Rath-Wiggins, Ph.D. is a former Tow-Knight Center fellow and a co-founder and CEO of Vragments.

Anna Sarmina, Ph.D. is a lecturer at the Taras Shevchenko National University of Kyiv, specialised in propaganda and information warfare.

Florian Stalph, M.A. is a Research Associate in Journalism at the University of Passau, Germany.

Alessandro Tanasi is a security researcher and lead developer at Ghiro Project.

Aric Toler is the lead researcher and trainer for Eastern Europe/Eurasia at Bellingcat and the Lead Digital Researcher for Eurasia at the Digital Forensic Research Lab of the Atlantic Council.

List of Figures

Fig. 11.1	"Der Trump-O-Mat". The Visual Story about Trump	114
Fig. 11.2	The digital version of *Der Spiegel*	117
Fig. 11.3	"Rost". The Visual Story about an Italian steel factory	119
Fig. 13.1	Crowdsourced geolocation via Twitter and Google Street View	141
Fig. 13.2	Crowdsourced geolocation of advertising sites via Twitter	142
Fig. 16.1	Aspects of context verification in press photography	173



CHAPTER 1

Digital Investigative Journalism

Oliver Hahn and Florian Stalph

> Condensation trails are mainly water-based by-products of aircraft engines, painting long, cloud-like shapes in the sky. Adding toxic chemicals—ones that have devastating consequences for our health—allows weather modification, mind control and human population control, just to name a few options of geoengineering. This theory of so-called chemtrails is—of course—a hoax. While this might appear as a ludicrous conspiracy theory, it is a widespread belief. In 2016, the Cooperative Congressional Election Study, a national study conducted by Harvard University, found that between "~30 and ~40% of the general US public appear to subscribe to versions of the conspiracy theory" (Tingley and Wagner 2017, p. 5). The study also shows that the biggest part of conversations about chemtrails is conspirative and can predominantly be found on Twitter.

This episode showcases how emerging disruptive technologies such as the Internet and particularly social networks channel disinformation and misinformation into the public discourse and influence the public perception

O. Hahn (✉) · F. Stalph
Centre for Media and Communication, University of Passau,
Passau, Germany
e-mail: oliver.hahn@uni-passau.de

F. Stalph
e-mail: florian.stalph@uni-passau.de

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_1

of the world around us. Oftentimes, certain digital spheres are—just like their inherent messages—driven by partisan allegiance or propagandistic tendencies. Facilitated through an increasingly fractured post-digital society, such messages can be deliberately and precisely shared among communities of interest and injected into public communication. An investigation by Albright (2017), research director at the Tow Center for Digital Journalism at Columbia Journalism School, has revealed Cambridge Analytica's effort to geolocate American voters and combine this information with real-time emotional sentiment analyses for targeting citizens via social media. Using that approach, it appears easy to feed relevant information to conspirators or everyone else, either consolidating their beliefs or shaking their faith. Such logics of a new ecosystem of information dissemination and consumption require new kinds of tools and new profiles of journalists.

Also, the ontology of journalism has changed: New formats and methods have changed journalists' everyday work routines that both emerging professional as well as more experienced reporters have to cope with. These innovations are accompanied by various challenges: masses of information provided online and in social media oftentimes leave audiences overnewsed but underinformed. As every piece of information gets digitised, journalists need to incorporate new practices to access digital sources, encode new information configurations and extract meaning of novel information types. Each kind of information requires specific approaches; digital formats can be highly distinct, contain accurate and detailed information and metadata. Specialists with distinguished skillsets have always been on the verge of breaking new ground for journalism innovation. A binary distinction between digital and analogue journalism is becoming extinct. New emerging technologies of communication are now an essential part of journalism and expand persistent aspects of a journalist's repertoire such as face-to-face communication with sources and eye-witnesses.

Investigative Journalism is heavily affected by the digital transformation. As the fourth estate, journalists are "quasi-constitutional watchdogs acting on behalf of a society's citizens" (Harcup 2014, p. 109). It is arguable whether journalists and affiliated news outlets manage to fulfil this role, acting as lapdogs not watchdogs (Benjamin 2014). Certainly, investigative journalism demands a lot from journalists: In order to uncover corruption, mismanagement, abuse of power or shady money operations and fraud, journalists are likely to stir up a hornet's nest bearing risks for their career, their whole profession and their own safety. By its very nature, investigative

reporters encounter resistance as they are trying to reinstate transparency and balance for the public good and according to democratic values. This is by no means desirable by those who are investigated. As a result, journalists face obstacles and forces they have to fiercely overcome. Powers that normally fight those who try to unmask them, may be hiding behind institutions, creeping in governmental bodies, exploiting loopholes and sailing close to the wind, or running sophisticated cons. Hence, citizens are in dire need of advocates who call out these wrongdoings. Accordingly, sources holding incriminating evidence are hard to access. Only few organisations may be willing to publish or to provide such evidence; usually, such information is well kept behind closed doors.

The digitisation of information—be it proceedings, contracts, annual reports, demographic statistics, economic indicators or audiovisual material—bears enormous implications for investigative journalism. On the one hand, journalists are now equipped with new tools that allow them to run analyses of high computational efforts, work statistical analyses and present their findings with a high visual appeal to broad audiences via the Internet. On the other hand, digital information is encryptable, decentralised and extremely prone to manipulation and tampering. To keep up with these challenges, journalists from both legacy media as well as independent activist platforms developed software, techniques and mindsets to prolong investigative shoe-leather reporting in the digital era.

In 2010, *The Guardian, The New York Times* and *Der Spiegel* reported on the Afghan War documents leak unveiling unknown facts hidden in numbers. The publication of these reports can be regarded as the ascent of data journalism or, at least, as its inauguration into the repertoires of reputable media outlets. Until then, data journalism had been a niche phenomenon we could scarcely observe in mainstream media. However, evolving internet technologies initiated the emergence of easy-to-use and mostly free tools to refine, analyse and visualise statistical data, which—due to the ongoing digitalisation of information in the past decades—would provide a new and seemingly inexhaustible source for journalists.

An online survey by Garrison (2001) revealed that internet technologies such as search engines and database services made their way into US newsrooms, showing that the Internet has been becoming a dominant source for news collection since the late 1990s. Later, Holovaty (2006) published a manifesto for database journalism. He calls on newspapers

"to stop the story-centric worldview" and to embrace data in their stories postulating news stories built around data-driven interfaces with regard for a combination of both text and data.

Today, data journalism can very well be considered as incorporated into the editorial structures of legacy news outlets such as *The Guardian* or *The New York Times* and we see more and more small data units on a local and regional level. Furthermore, thanks to the mere existence of stand-alone data-driven explanatory and analytical data journalism sites such as Vox, FiveThirtyEight or Quartz, and thanks to the inclusion of data journalism within journalism education and the growing consideration by the scientific community, data journalism is no more on the fringes but a well-known and widely discussed journalistic phenomenon.

Data journalism is symptomatic for the data age: Journalists need tools to report information that is now saved in digital forms. But not only words and statements are abstracted and are turned into spreadsheets, and not only demographic or economic statistics are stored as intelligible datasets. Audiovisual materials such as photographs and videos are other information mediums that challenge digital journalism. It is particularly difficult to authenticate visual and audiovisual material and to identify the originators of those contents. How can journalists retrieve data, pictures and voices that can enlighten stories, provide background information or will result in exclusive scoops? Data-driven journalism and its techniques of analysing and visualising statistical big data and new visual-analytical tools are cutting-edge approaches to uncover the truth and tell the story behind the story; whereas the former has found its way into most newsrooms, visual analytics techniques are predominantly applied by independent specialists, experts working for independent platforms or institutions affiliated with academia. This edited volume displays various approaches, technologies and projects that employ innovative practices within their investigations. By no means does this outset imply that new technologies and practices for investigative newsgathering replace traditional techniques and make investigations exclusively desk-bound; however, new approaches certainly do enhance access to a variety of unique sources and voices.

This book compiles chapters by both practitioners and academics. Contributions range from engaging anecdotes or authoritative instructions to contextualised discussions. As a result, chapters differ regarding their format and intent. If you are a practitioner yourself, some chapters hopefully prove to be useful hands-on manuals while other chapters

provide additional insights at a meta-level. At the same time, we hope to give our academic readers exclusive insights into applications of modern techniques while embedding these cases in the broader context of journalism innovation. When structuring this volume, we had in mind that some of our readers will find certain chapters more personally valuable than others. Therefore, we arranged the contributions in three parts, focusing on three central pillars of digital investigative journalism.[1]

The opening part introduces data journalism and computational journalism and other data-based journalistic approaches. It features international data journalism pioneers who stand out due to their influential contributions. This part will be practitioner-oriented. Reputable journalists share their exclusive insights into data-driven investigations carried out by e.g. *The New York Times*, the International Consortium of Investigative Journalists (ICIJ), or the Organized Crime and Corruption Reporting Project (OCCRP). This part shows how data journalism was incorporated as an investigative approach, the potential of data for journalists and how to find stories in datasets.

The following part introduces experimental and progressive approaches of journalistic investigation. Journalists and technologists working for Bellingcat describe highly innovative visual analytics approaches in order to examine digital information and to build evidence that would never be retrievable otherwise. This part heavily focuses on visual information and visual storytelling. The central aspect is how digital technology can be used for visual analyses, unveiling hidden stories and visualising scenarios with the help of computer technology. By merging computer forensics with journalistic inquiries, characteristics of digital contents can provide insights into scenarios that were previously inaccessible.

The closing part discusses the implications of digital techniques for journalism and how digital spheres change public communication and media production. By directly expanding on the previous parts, it will focus on methods to identify originators of social media content and to check authenticity of such questionable material by examining practices introduced throughout the book within societal contexts. Particularly information disseminated via social media can be unique and highly relevant evidence; however, when using this kind of information, journalists

[1]We thank our research assistant David Liewehr, who sadly passed away before he could see this book completed, for his support in editing this volume.

might run the risk of falling into a trap of incorrect or even manipulated content. Against this background, this part also focuses on propaganda as well as on fact-checking, and debates (mis)information flows. There will be a further discussion regarding need for verification, as social media content can easily misguide journalists. In this part, chapters spotlight different challenges of social media content as well as official data that may also be prone to manipulation and misinformation; at the same time, journalistic investigations need to adapt to certain characteristics of such information. Besides practitioners who illustrate their means of information verification, academic essays construe surrounding aspects of misinformation and defective news proliferation in post-truth spheres.

References

Albright, J. (2017, October 13). *Cambridge Analytica: The Geotargeting and Emotional Data Mining Scripts* [Web log post]. Retrieved from https://medium.com/tow-center/cambridge-analytica-the-geotargeting-and-emotional-data-mining-scripts-bcc3c428d77f.

Benjamin, M. (2014). Why Are the Media Playing Lapdog and Not Watchdog—Again—On War in Iraq? *The Guardian*. Retrieved from www.theguardian.com.

Garrison, B. (2001). Diffusion of Online Information Technologies in Newspaper Newsrooms. *Journalism, 2*(2), 221–239. https://doi.org/10.1177/146488490100200206.

Harcup, T. (2014). *A Dictionary of Journalism*. Oxford: Oxford University Press.

Holovaty, A. (2006). A Fundamental Way Newspaper Sites Need to Change. *Adrian Holovaty*. Retrieved from www.holovaty.com.

Tingley, D., & Wagner, G. (2017). Solar Geoengineering and the Chemtrails Conspiracy on Social Media. *Palgrave Communications 3*. https://doi.org/10.1057/s41599-017-0014-3.

PART I

Data-Driven Investigation and Computational Journalism

Structured information is omnipresent. Every piece of information, every statement made by politicians or entrepreneurs, every bank transfer, every election result, every train arrival, every deadly bullet fired—they are all stored in databases in one way or another. Within those databases lie datasets containing information of elusive value for both, for those who collect and commercialise them, and for those who devote themselves to produce and publish reports based on data on behalf of the public. Over the past decades, quantifiable data has been becoming increasingly popular: Philip Meyer's so-called precision journalism has made popular to employ social sciences methodology in journalism. At the same time, newsroom staff has started computational analyses of statistical data. The journalistic use of data has peaked once widespread access to the Internet has become available. This has paved the way for the growth of platforms and forums that offer tools for data analysis, interactive visualisation, dissemination and collaboration.

Following the so-called computer-assisted reporters, new generations of 'hybrid' journalists were born: e.g. journo-coders, data journalists, programmer-journalists. They all have in common to combine journalism practice with methodologies drawn from sciences and innovative technologies. This has sparked heated debates within the media industry and also in academia. Data and computational journalism seem to become the new credo. Over the last couple of years, these journalistic currents have been installed in traditional newsrooms, small local outlets, in independent or niche journalism platforms, and in journalistic tech start-ups.

Check for updates

CHAPTER 2

The Rise of Investigative Data Journalism

Andrew W. Lehren

My journey into data journalism did not begin with some massive database. It was a long way off before I would get involved in combing through secret diplomatic cables or the Snowden trove of spy documents. It was before I really appreciated how statistics could further my journalism, to the point where it would be using statistics that would lead to the recall of a quarter of a million Nissan Altimas because its airbags were blinding passengers. The arrests of those involved in defrauding a major railway's disability pension system. Or spark the Pentagon to overhaul the way it delivers medical care to its soldiers and their families. It was before I learned about computerized mapping, and how it could help uncover policing that disproportionally targets African Americans, or how real estate developers reshape flood maps that later haunt unsuspecting property owners when hurricanes strike.

I was at a small weekly newspaper in Philadelphia, covering City Hall and trying to find enterprise stories that the bigger media organizations were missing. I sometimes got a good tip. But those often went to the veteran reporters at other publications. A friend told me about Investigative Reporters & Editors, a non-profit in the United States that helps journalists learn how to dig. I went to a conference and learned

A. W. Lehren (✉)
NBC News, New York, NY, USA

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_2

how to do a better job combing through public records. I started checking out court filings more often, became sharper going through business documents and began examining campaign reports. That was a start.

Poring through an endless pile of dusty documents can be a slog. Piles of paper. What would keep me going would be the results. The questionable economic deals that would be spiked because of what I found. The massive government contract that would be rescinded. The politician who claimed his prowess with finances only to be shown to be not everything he advertised.

Then other journalists started saying that if I learned how to use a spreadsheet, I could begin to do a better job going through public records. If I learned how to use a database manager, I could go bigger. I paid my way to go to some of the first seminars for journalists in the United States to see if this would live up to its promise. I became hooked.

If you wonder why I mention going through documents, it is because that is the image I have kept in my mind in the more than two decades I have been doing data journalism. It contains fundamental points that I sometimes worry about with the facile way information is talked about sometimes these days.

The documents I pored through would sometimes be incomplete. Pages could be inexplicably missing. Or someone filled out government forms incorrectly. Or to really understand an inspection report, you needed to learn the codes mysteriously noted in a corner. You needed to understand exactly why the documents were collected, to begin with, and what might be hidden from public inspection.

You also needed to know these were not an end, but a start. They sparked more questions. They altered the way you would do interviews. And who you asked questions. More importantly, you had to keep in mind that a pile of documents was fundamentally boring to most readers. Your job was to find the story—the important journalism—and find an engaging way to incorporate it in the telling.

At one point at *The New York Times*, I was feeling a bit burnt out. I had finished contributing to one project that won a Pulitzer. Run ragged helping cover the massive BP oil spill in the Gulf of Mexico. And worked on several enterprise stories on US casualties in its wars in Iraq and Afghanistan.

I asked my editors if I could spend a couple weeks just writing some stories for sports. Just for a change of scenery. I have been a long-time

runner. I had some half-baked ideas about marathons. I thought I could use a brief busman's holiday before diving back re-energized into weightier subjects.

But the last thing I wanted to do was a story that someone else had already done. I would leave the heart-warming feature about a runner who overcame obstacles for another day. I started casting around. What public records even existed about running? I thought of the phone apps that friends and family use to track loved ones in races. I called the organizers of the New York City Marathon to see if I could get the underlying data from the most recent race. It turned out to be the same data that newspapers and websites use to list results. Pretty boring stuff.

I put it into a database manager. About 50,000 records. And started asking questions to see if anything would pop up. For those unfamiliar with marathons, this is how it works in New York. You start off in Staten Island. Cross the arching Verrazano Bridge into Brooklyn. Cut straight into Queens. Take the 59th Street Bridge into Manhattan, where runners are greeted by throngs of cheering crowds. Turn up to the Bronx. Cut back down through Harlem and into Central Park. You finish having covered all five boroughs 26.2 miles. In one of the world's biggest running races.

What I could not figure out was why I could follow most runners throughout the race. But not all of them. The race organizers have computer chips for all the participants, no matter how slow. They could track them every few miles. I had times for almost everyone throughout the race. But not everyone. It could not be a faulty computer chip. The chips worked at the beginning and end of the races. But for five or more split points in a row along the course, there were no times.

I began wondering about other documents that are public. One feature of every marathon is the souvenir photograph. Get a picture to commemorate your accomplishment. And the photo sites are public. Of course. What better way to make it easy for you or your family to find that image. I started using it to check the runners with five or more missing split times. No images of them along those stretches of the course either.

Now I had some interesting questions for the race organizers. They quickly conceded: Welcome to the dirty underbelly of large marathons. Dozens and dozens of people cheat every time. Cheaters hop on subways in Brooklyn. Or bleed into the crowd greeting them when they enter

Manhattan. And then magically re-emerge in Central Park to collect their medals.

I had not seen that story before. This took me about a day or two to puzzle out. I could probably pitch this for the front page. I checked the archives of various publications. There had been occasional stories about the odd cheater here and there. But nothing about the scale. The sports editor was intrigued.

I suppose I could have filed something in a few days. Track down some of the cheaters. I had their names. Amusing. But I paused and asked, what's the harm? What's the foul? So some people feel compelled to cut corners and collect a finisher's medal they did not earn. The overall winners were not affected. I paused and started re-evaluating results. Then I realized that in subcategories, like the fastest man in the aged 60–65 category, the winners were being affected. I tracked down those involved.

When I reached Alan Miller, who should have known on race day that he won his age grouping for older men, he asked how I even figured out his story. I told him. He said, Listen, I've run the marathon 25 times. I get ready for it with my buddies from the New York Police Department. Afterwards, we all go out for drinks at Micky Mantles. One of the landmark waterholes then in New York. If I had known I had won that day, it would have been the biggest day of my life. The drinks would have been on me. I would have celebrated with my friends like there was no tomorrow. And now I have cancer. I may never run another marathon.

This is why I will comb through data, no matter how boring. To find the stories of the Alan Millers of the world. To find bigger patterns that other journalists may miss.

It is this mindset that I had used about a decade earlier working for NBC News.

A young unarmed black male, Timothy Thomas, had just been shot and killed by Cincinnati police. Riots broke out in the streets as African Americans protested what they argued was a criminal justice system that consistently worked against them unfairly. Front page stories followed the demonstrators. The protests dominated cable and evening news broadcasts.

The president of the news division asked me and another producer to spend a few weeks, and see if we could find a story others had missed. I watched, again and again, the dash camera footage of the officer in the minutes leading up to the shooting. All the officer really knew about

Thomas was a description and that he had 14 outstanding warrants. For a 19-year-old, that was a lot. Sounded like a hardened criminal. That was the impression in much of the coverage.

I thought I would check the court records while others covered the protests. It turned out every one of the warrants stemmed from a traffic stop. More than that, they were not typical citations. Not for speeding or running stop signs. They were for small things an officer would have a hard time knowing before he pulled over Thomas, like not wearing a seat belt and not having proper paperwork.

Was this unusual? I used Ohio's open records laws to get data from the Cincinnati police. When it came to moving violations like speeding, blacks and whites were pulled over at about the same rate. But when it came to these non-moving violations, blacks got these tickets three times as often as whites.

This took about two weeks to sort out. We told our boss. He said that sounded promising. Now see if this pattern held true in cities across America. A fantastic assistant, Ben Vient, helped me file open records requests for more than 15 cities we knew were required to keep similar data. Almost everywhere, the patterns held up.

Criminal justice experts told us the key ratios to examine. When police said disparate ticketing happened only in high-crime neighbourhoods, we mapped the data. That was not true. Blacks and whites were treated differently regardless of neighbourhood crime rates.

The result, produced with Jason Samuels and hosted by John Larson, was an hour-long documentary that was the most-watched programme that evening in America. It went on to win a slew of awards. More important, it raised compelling questions about policing.

In the wake of the Ferguson, Missouri riots in 2014, I revisited this work with Sharon LaFraniere at *The New York Times*. Sadly, the patterns still held true. We also delved deeper into statistics on searches, and found police in many places conducted speculative searches far more often on African American drivers. Particularly since the searches were supposed to uncover guns or drugs. The statistics showed that in most places, police were far better on average when they searched whites. When they searched blacks, which they did far more often, they disproportionally failed to discover contraband. Our story was centred on North Carolina, and the resulting front-page story led to a storm of public hearings about policing there.

Beyond database work, the statistics and mapping were important. Each discipline has its own key variables. Like preparing for any interview with a public official, knowing the right measures to check can lead to most revealing answers.

Several years earlier, when Sharon and I were combing through lawsuits that the United States lost in federal court, we found disturbing anecdotal cases involving military hospitals. There are 55 such hospitals around the world, set up to help more than 9 million service members and their families. We wanted to know the big picture. We asked the Pentagon for data and statistics. Every major civilian hospital chain maintains such metrics. The military countered that it kept no such information.

We fought a Freedom of Information Act request for more than a year. Eventually, we started getting data the military has said it did not maintain. When the Army surgeon general, who oversees all its hospitals, told us his facilities had no problems, we pointed out several with troubling surgical complication figures. Shortly the head of one of the hospitals lost that position. We learned about the importance of shoulder dystocia in birthing; how it is a marker for troubled deliveries. Checking those figures pointed out a disturbing pattern particularly among the military's smaller hospitals. The resulting series of stories led then US Defence Secretary Chuck Hagel to revamp oversight of military hospitals.

That was not my first time closely examining military numbers. Jim Dao and I were trying to turn around a quick look for *The New York Times* on military suicides when we noticed numbers not adding up. The published rates, and the testimony by military leaders before Congress, compared with overall suicide figures, did not make sense. We paused and dug deeper. It turned out the Pentagon had been systematically underestimating its suicide rate. It was not using established methods deployed by the US Centers for Disease Control, the touchstone for such calculations. This meant the Pentagon missed for years the rising suicides happening during the Afghanistan and Iraq wars, and if it was measuring these properly, it would have known it had a problem and perhaps instituted measures that could have saved some lives. Even more, it had neglected to even track suicides involving reservists and others. The Pentagon, in the wake of the stories, revamped its tracking of suicides.

I should pause here and be clear. I am just one of the hundreds of data journalists now doing this kind of work in the United States and around the world. If anything I write can encourage more to enter this field, then the writing of this chapter is worthwhile. I know there are plenty doing incredible work and do not mean to put mine ahead of any others, only to say that, if anything, it is part of a larger mosaic. Also, there are visionary journalists before me who did groundbreaking work well before me. There was Phil Meyer, who won a Pulitzer examining the forces behind the race riots in Detroit in the 1960s. There were Barlett and Steele, two Philadelphia Inquirer reporters who looked at how judges were meting out sentences in the 1970s. They used computers, but to feed in data, they had to use punch cards stacked for old mainframe machines.

There were nurturing stalwarts like Pat Stith, a North Carolina journalist who helped IRE organize its first computer-assisted reporting conference. Stith would go on to win his Pulitzer exposing how corporate hog farmers had hijacked state oversight and were damaging the environment.

Sometimes journalists can pause and wonder if using some new technology is a proper way to do reporting. The debate happened more than a century ago when the telephone was introduced. Knowing how to cull data is just one more tool.

Journalists can approach data differently than those more trained in computer sciences. Take, for instance, matching databases. Traditional IT managers compare data sets that were designed to talk with each other. Journalists may wonder if the payroll list of school teachers includes registered sex offenders.

When *The New York Times*, along with *The Guardian* and *Der Spiegel*, was given access to the Iraq War Logs funnelled by Chelsea Manning to Wikileaks, I joined the team of reporters going through the data. Unlike the later trove of diplomatic cables, these were pocked with military abbreviations and jargon. Some reporters went through record by record, looking for a big find.

Knowing database tools led us to check the completeness of the files. Unlike the earlier Afghanistan war logs, we could find almost every incident involving a US casualty. That was reassuring, particularly since we could not expect the Pentagon to vouch for the completeness of a purloined database. Jim Glanz and I wondered about violent incidents involving private security contractors. Most readers knew about

a particularly controversial firefight involving Blackwater. Jim, a former Baghdad bureau chief for *The Times*, long suspected there were more incidents than officials acknowledged.

I stepped away from the secret room where we had sequestered the data and began going through public records. Audits, contracts, hearing transcripts, news accounts. I assembled a list of more than 100 private security contractors that were known to be involved in Iraq. I took that database and flung it against the text inside the war logs. I then ran more queries against the war logs to see if we missed others. Jim and I were then able to stitch together a front-page story about how indeed there were far more incidents than the public had ever been made aware of. To tell the story, we mined social media. Barbara Gray, an ace researcher at *The Times*, trolled LinkedIn resumes to find contractors that had been present at the incidents we now could document. That helped confirm what we were seeing, and helped provide a more compelling narrative that the staccato abbreviations in so much of the data.

We took similar approaches with the diplomatic cables, stitching together from seemingly unconnected passages how companies around the world had helped North Korea's missile programme and provided precursor ingredients for Syria's chemical weapons programme. Seemingly unconnected records helped us document how unregulated tainted Chinese chemicals were becoming part of the world's brand-name pharmaceuticals, sometimes with deadly consequences.

You may have noticed a pattern. Because data journalism can be time-consuming and complicated, you can get perhaps the best results when you work with colleagues to share the burden. I cannot acknowledge enough all the great reporters I have had the chance to work with on stories like these. More minds may lead to better questions for the data, and then stronger stories. As more newsrooms collaborate on data-driven stories, the depth of the reporting only increases.

When I was running IRE's database operations in the mid-1990s, I saw the benefits of many reporters examining the same data. That small staff included five journalists who would go on to be involved in Pulitzer winning journalism.

Another pattern: the tools keep evolving. Mining social media is becoming ever more important. Programmes that can help tease out connections in social networks, and can help report on people, companies and governments are playing a larger role. I recently rejoined NBC News as a senior editor on its investigative team. The organization now

has several journalists whose sole job is to comb through social media and find information. The tools and methods are only expanding.

Data journalism can be a great leveller. It helped me find stories way back in Philadelphia that my more established rivals had not found. I teach a class at a New York journalism school, and I have seen my investigative reporting students find big-picture patterns in data for stories that their more seasoned competitors overlooked. I look now in awe at the international growth in these skills transforming journalism in countries across the world. The Panama Papers is evidence of that. So is the work at the Global Investigative Journalism Network and so many other investigative reporting centres spanning the world.

The types of tools out there keep increasing. If it ever seems daunting, I think back to that small newsroom in Philadelphia, and remember: It is all about finding stories that matter.

Check for updates

CHAPTER 3

Zeroes and Ones: Investigating with Data

Paul Bradshaw

Introduction

When a source of information is rich, and important—and growing—it would be foolish to ignore it. Just as it seems strange now to write an article about a war without speaking to those affected (as was once common practice), it may well become increasingly unusual to write about sport, fashion, business, the environment, or politics, without some sort of understanding of the rich seam of structured information that underpins the decisions that are made in that sector, provides context to its history and trends, and helps identify the changes—the *news*—that it faces and has yet to address.

Investigative journalism in particular has increasingly come to rely on the opportunities that data journalism techniques open up. After all, many of the qualities of investigative journalism—dogged persistence, identifying systemic problems (patterns) and shining a spotlight on what was previously hidden—suit data journalism well. Data journalism techniques have helped open investigative opportunities to a much wider range of journalists than just those who have "investigative" in their job title, while *data itself*—whether used to inform decision-making, or in the form of algorithms that shape what we can and cannot do—has

P. Bradshaw (✉)
Birmingham City University, Birmingham, UK

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_3

become a form of power which investigative journalists seek to hold to account.

In this chapter, then, I will outline that range of sources that journalists use to get data, how we find the stories in that data and tell them—and what the pitfalls are that we seek to avoid along the way.

Which Zeroes and Ones? Finding the Numbers

The widespread digitisation of our lives, of power and its operations, and of almost every sector of industry and culture, has a useful side effect for journalists: quantification. From notes in music and colours in images, to school absenteeism and political lobbying—almost everything can be measured, and typically is. The challenge is not the existence of numbers—but rather getting hold of those numbers, in a condition that makes them open to investigation.

In particular, the following qualities have a significant effect on how accessible a data set is to a journalist:

- The **machine readability** of data: Is it in a format which can be immediately analysed, such as a spreadsheet—or is it instead in "closed" formats like PDFs and images?
- The **concentration** of data: Is it available from one location, or scattered across dozens of websites?
- The **explicitness** of data: Is it immediately obvious that this is structured information, or does the journalist have to exercise creativity to perceive the opportunities for data journalism?
- The **predictability** of data: Does the information always obey a particular pattern, or is it inconsistently stored/presented?
- The **quality and granularity** of data: Is it detailed or sketchy? Aggregate or granular? Does it need cleaning?
- The **availability** of data: is the data "pushed out" to reporters, or do they need to request it using legal means such as Freedom of Information? Is it published in an obvious place, or hidden away in a way that makes it difficult to discover?

The way that those factors combine determines the difficulty of doing journalism with that data. Human-readable data needs to be converted in some way; scattered data needs to be collected; dirty data needs to be

cleaned; and implicit data not only requires imagination to be able to see it *as* data, but the technical ability to collate and structure it.

When starting out in data journalism it is natural to look for sources where the data is ready-made for analysis, in the form of spreadsheets. This is a practical decision, but also one to consider critically, because such practical decisions also shape the types of stories that we are more likely to do as journalists—and the types that we are unlikely to report.

Open data initiatives, for example, have expanded significantly over the last decade, but their implementation has often been criticised. One report on the use of open data by journalists noted that it was "frequently out of date, incomplete, or of limited new value" (Stoneman 2015), while in the US The Sunlight Foundation's Ellen Miller noted that "The Department of the Interior seems to think that the population count of wild horses … [is important enough to publish] but records of safety violations … are not" (O'Reilly 2010). Control is an issue, too: when Donald Trump took office, the publishing of White House visitor logs was discontinued and the Environmental Protection Agency announced that it would close its open data website. Although a backlash prompted the agency to claim the announcement was a mistake (Gallucci 2017), the point had been made.

Open data comes also almost entirely from the public sector, which in many countries is also subject to access to information laws. As a result, it is subject to much greater scrutiny from journalists who use data raising the question of whether it allows other sectors to escape similar scrutiny. Private companies, for example, increasingly operate public services, exercise significant power over people's lives, and are central to some of the biggest stories of the age, from tax avoidance and environmental change to inequality and insecure employment, but these typically make much less data publicly available.

Many reporters will try to understand the systems that operate in the field that they are planning to cover. There may be charities and non-governmental organisations (NGOs) who have to regularly deal with problems in that field—they may have some sort of data, even if they are not aware of it (for example, numbers of calls received or locations that they have visited). Businesses in the field will have to be registered, and their directors and shareholders may be named (OpenCorporates has data from dozens of company registers around the world). There may be annual reports and mechanisms for recording and monitoring information on what organisations do—this might be

requested, or it might be leaked. Jobs and contracts will be advertised online, social media accounts will be updated.

In fields that inspire passion, there may be fans collecting information, too—sport is a particularly good example where you can find lots of data published by devotees and participants. Music, art, and fashion have this characteristic, too.

It is often worth thinking about technology services which might have an API for the data you need—an API (Application Programming Interface) is a way of "querying" data held by an organisation. For example, in one story about unsolved crimes, and how that varied between forces (Wainwright and Bradshaw 2017), I used the UK police force API to gather 37 million records.

Data Leads and News Values

Journalism has a curious, misleading language: we talk about "finding stories", when really what we mean is that we find *leads* for stories—those qualities that we have learned, over time, distinguish something newsworthy from something ordinary. There is a large body of research around how journalists make that distinction, and turn leads into stories, most of which make some reference to Galtung and Ruge's seminal *news values* study of 1965 which identified a number of factors broadly fitting into three categories: impact, audience, and practicality (Galtung and Ruge 1965).

Harcup and O'Neill's (2016) study attempts to bring things up to date by identifying how news values play a major role within modern journalism (Harcup and O'Neill 2016), and these can all be related to data journalism, too. "*Bad news*", for example, is one of the first things a journalist might look for in a given data set (Harcup and O'Neill 2016): If there are a large number of mistakes being made, or a lack of action, then that might be an important story to tell. Sorting the data to show which entries do worst on one particular measure can quickly take us to one potential story.

But often the numbers on their own are not meaningful: It is important to establish the context that makes those numbers significant. One source of context is time: Are those numbers getting worse or better? In order to establish that, journalists must calculate the difference between different time periods (years, preferably, to avoid seasonal variation due to external factors like weather).

Another source of context is the promises and claims made by people in power: have they claimed that something is not a problem when it is? Have they promised to tackle something but failed to do so? This also satisfies another news value: the involvement of the "*power elite*" (Harcup and O'Neill 2016). And fact-checking the statements of that power elite forms a significant category of data journalism in its own right.

The *agenda* of an organisation plays a role here: "Bad news" which undermines that agenda may be unlikely to be given the resources it needs, while stories which support it may be positioned prominently and promoted heavily: disability rights, for example, fit BuzzFeed's agenda strongly enough for it to devote a week to investigations related to the topic in 2017, while The Daily Mail chooses to run with "Workshy map of Britain revealed: Thousands of incapacity benefit claimants found to be capable of working" (Williams 2013).

Conversely, "*good news*" is also seen as an important news value (Harcup and O'Neill 2016). Within data journalism this can take the form of *solutions journalism*: stories which sort data to find the entries with the most positive values, and what might be learned from those bodies (including if the data reflects the reality). Research has suggested that readers engage more with stories using such an approach, but also feel more empowered to act themselves (Curry et al. 2016).

"*Magnitude*", a fourth news value, relates not just to the sheer scale of some data stories, but also their political impact (Harcup and O'Neill 2016). When the Bureau of Investigative Journalism set out to establish the number of mega farms in operation in the UK, it led directly to questions in Parliament (Davies 2017). But stories involving large amounts of data are inherently interesting, too. "Megaleaks" from LuxLeaks to the Panama Papers are striking in scale, if often technically challenging: Querying that amount of data requires coding and a considerable amount of planning and coordination.

We look for "*surprise*" in data in the form of change (Harcup and O'Neill 2016): numbers rising or falling, ideally unexpectedly. But unexpectedness might lie in something that we hadn't thought to look at in a certain way before. The Mirror's story on "Which singer has the best vocal range in the UK - No, it's not who you think" (Beech and Scott 2014) uses data to challenge what we think we know about popular music (it also meets two other news values: "*celebrity*" and "*entertainment*"), while The Seattle Times surprises us by showing the role of

women in the art world in connections between mentors, collaborators, and peers in the scene (Romano 2017). Sometimes the surprise lies in the storytelling: the New York Times's "*You Draw It*" series of stories involving interactive charts invite the readers to test their assumptions about political and social trends.

"*Drama*" and "*conflict*" may not sound like data journalism values, but when organisations refuse to disclose data or admit that they don't collect it, it sometimes leads to a story about "criticisms" and "attacks" on that organisation for attempting to dodge accountability (Harcup and O'Neill 2016). There may even be a "fight" for the data to report, too. When the BBC Data Unit identified that one in five MPs were still employing members of their family despite new MPs not being allowed to do so, the story led on "Give MPs deadline on hiring relatives, campaigners urge", rather than the figure itself (Sherlock and Bradshaw 2017).

The importance of "*exclusivity*" as a news value often leads data journalists to seek out data which is not easily accessible (Harcup and O'Neill 2016): The importance of scraping skills (automating the mass collecting of data from multiple webpages or documents), drafting successful FOI (Freedom of Information) requests, or the use of coding to analyse large data sets can also contribute to the strength of a story and the ability to say "No one else can do this story like we can". BuzzFeed's investigation into suspect betting patterns in tennis, for example, included an article devoted to "How BuzzFeed News Used Betting Data To Investigate Match-Fixing In Tennis" (Templon 2016).

Notably, Harcup and O'Neill suggest that "arresting *audio-visuals* [emphasis added] are certainly worth listing as a news value in their own right" and that "the findings point towards the importance of a concept we call *'shareability' [emphasis added]*" (2016). And while not all data journalism employs data visualisation, its role is demonstrated by the inclusion in many data teams of individuals with some form of design skills. Some data stories are entirely or primarily visual, from the "home of Telegraph data journalism" TeleGraphs (telegraph.co.uk/news/data), to The Wall Street Journal's interactive graphic "Donald Trump, His Children, and 500+ Potential Conflicts of Interest" (Wu 2017).

Visuals are just one element of shareability—interactivity and personalisation is another. The Guardian's investigation into "the insidious creep of pseudo-public space in London" invites users to "Explore the map – and tell us what we're missing" (Shenker 2017).

The potential for interactivity is increasingly a consideration when looking at potential story ideas.

Relevance to your particular audience is one of the most crucial elements in choosing and pursuing stories. An international story might be important, but if your audience is local, it may not be worth dedicating resources to—and vice versa. The Trinity Mirror data team, for example, looks for investigation ideas and data sets that can provide multiple stories for its different local and national publications (such as crime, schools or hospitals), a model which has also been adopted by a number of other organisations.

There is also one final news value—the "*follow-up*"—that relates to how we get stories (Harcup and O'Neill 2016). This is the subject of the next section.

Story Ideas: From Follow-Ups to Proactive Stories

Traditionally reporters operate in either a reactive or proactive mode. Reactive stories see the journalist responding to an external event of some sort: it might be a government announcement or a press release; a disaster or a protest; the publication of new research or something that someone has said on Twitter. Follow-ups are a common reactive format—especially in journalism.

Proactive stories, in contrast, originate from the journalist or newsroom in some way: it might be an idea to interview someone interesting, or to find out how something works; it might be a tip-off about something newsworthy or a nagging question that the reporter has.

Data journalism stories can also be reactive or proactive: the typical reactive data journalism story, for example, sees the reporter responding to some new official statistics, but they can also be led by the broader news agenda to establish a "bigger picture".

Proactive data journalism stories tend to set the news agenda rather than follow it—for that reason investigative stories tend to be proactive by default. BuzzFeed's investigation into evidence of match-fixing in tennis, for example, involved devising an algorithm to analyse gambling on professional tennis matches over the past seven years (Blake and Templon 2016). And The Las Vegas Sun's Do No Harm series (lasvegassun.com/hospital-care) used data requests to find out how many preventable injuries, infections and surgical mistakes there were in hospitals.

The distinctions are not always clear: the Telegraph's MPs' expenses stories of 2010, for example, were in many ways "reactive" because reporters were reacting to a data set that they had acquired, rather than beginning with an idea for a story (Winnett and Rayner 2009). However, the data set was only available because three journalists had used the FOI Act to request those expenses—and then made their case in the courts.

Proactive stories typically begin with a hunch—a hypothesis or hypotheses—that the journalist suspects may prove to be true when tested against the evidence. This can be useful in focusing an investigation, but can also mean overlooking information which may lead to more important stories. However, data-driven techniques are beginning to offer alternatives to this approach: in their research on data journalism in a Chicago newsroom, for example, Parasie and Dagiral identify in data-driven methods a "challenge to [the hypothesis-driven] epistemology" (Parasie and Dagiral 2013): reporting that begins with a data set and looks for unusual patterns, while "data analysis begins with discovery rather than a hypothesis" at the Wall Street Journal: team leader Jon Keegan notes that "Sometimes when we visualize the data, it can become the thesis of the story" (Usher 2016, pp. 118–119).

Legal and Ethical Dilemmas

Data journalism is bound by the same ethical and legal codes and considerations that the journalism profession as a whole chooses to abide by. The requirement to be accurate and to be fair, are key foundations: Data should be treated like any other source, and a journalist working with data should seek to check that it is accurate (through accessing a second source, for example) and that they have attempted to contact all parties involved. Sometimes conversations can highlight incorrect assumptions about the data, or flaws that the reporter was not aware of.

Surveys and polls can be a particular pitfall for reporters—their appearance of scientific rigour is often exploited by PR officers to obtain free publicity for their clients. However, journalists should remain sceptical—particularly about surveys which do not disclose their sample size or methodology. An understanding of statistics, and a willingness to report the margin of error, is important for journalists dealing with surveys. Editorial guidelines on reporting polls, such as the BBC's (bbc.co.uk/editorialguidelines/guidelines/politics), might also be consulted.

When stories involve publishing large amounts of data points which would be impossible to verify individually, the reporter must make clear where the information came from—and if possible offer a route for corrections to be made.

Stories involving data can be libellous like any other, so the journalist must be confident that reporting is not misleading, and that they have the evidence to back up any assertions. And publishing data can constitute a breach of data copyright law or database rights, too (Bradshaw 2014).

The protection of sources is becoming an increasingly important ethical consideration for journalists. News organisations could once mount a legal defence against revealing their sources in court, but now individual journalists must be able to plan a technical defence first—otherwise, sources may be compromised without either party being aware (Bradshaw 2017). Threat modelling is an important first step on this front. The data that sources supply—which may be hacked or in breach of their employment contract—may also raise legal and ethical considerations.

Many ethical dimensions are not a question of the "right" way to do things, but rather an ability to negotiate the tensions between two competing ethical demands. Within data journalism such a tension often plays out between a person's right to privacy and the wider public interest: when dealing with sensitive data, journalists might decide that one overrides another, but there is rarely a perfect decision.

Conclusion

Data journalism techniques offer enormous possibilities for journalists looking to investigate stories in the public interest: in the newsgathering stage, they make it possible to gather amounts of information that would not have been feasible for most journalists two decades ago, lowering the barrier to exclusive reporting. In terms of analysis, they allow the journalist to establish the magnitude of an issue that might otherwise be visible only through victims' anecdotes. And in terms of storytelling we can now create experiences that are social, visual, personalised, and interactive—if we want to.

At every stage of the process, however, the data journalist must be aware of the choices that they are making and the reasons for those. Are we choosing data because there is a story that needs telling—or choosing

stories because the data is easy to find? Are we treating data with the same scepticism as we would treat a human source? Does a fetishisation of numbers lead us to ignore the human element in a story? And are we telling it in ways that have the biggest impact with the widest range of people?

As we tell more stories, gain more experience in making those choices, and share that with others, we can only hope that we will be better prepared to make the right ones.

References

Beech, R., & Scott, P. (2014, September 26). Which Singer Has the Best Vocal Range in the UK—No, It's Not Who You Think. *The Mirror*. http://www.mirror.co.uk/news/uk-news/singer-best-vocal-range-uk-4323076.

Blake, H., & Templon, J. (2016, January 17). The Tennis Racket. *BuzzFeed*. https://www.buzzfeed.com/heidiblake/the-tennis-racket.

Bradshaw, P. (2014). Ethics in Data Journalism. In L. Zion & D. Craig (Eds.), *Ethics for Digital Journalists: Emerging Best Practices* (pp. 202–219). New York: Routledge.

Bradshaw, P. (2017). Chilling Effect: Regional Journalists' Source Protection and Information Security Practice in the Wake of the Snowden and Regulation of Investigatory Powers Act (RIPA) Revelations. *Digital Journalism, 5*(3), 334–352. https://doi.org/10.1080/21670811.2016.1251329.

Curry, A., Stroud, N. J., & McGregor, S. (2016). *Solutions Journalism & News Engagement*. Retrieved from Center for Media Engagement Website: https://engagingnewsproject.org/wp-content/uploads/2016/03/ENP-Solutions-Journalism-News-Engagement.pdf.

Davies, M. (2017, July 21). US-Style Livestock Farming Will Not Come to the UK, Says Michael Gove. *The Bureau of Investigative Journalism*. https://www.thebureauinvestigates.com/stories/2017-07-21/michael-gove-intensive-farming-defra.

Gallucci, M. (2017, April 24). The EPA Won't Be Shutting Down Its Open Data Website After All. *Mashable*. http://mashable.com/2017/04/24/epa-open-data-website-shutdown/#YO0IkTowpaqj.

Galtung, J., & Ruge, M. H. (1965). The Structure of Foreign News: The Presentation of the Congo, Cuba and Cyprus Crises in Four Norwegian Newspapers. *Journal of Peace Research, 2*(1), 64–90. https://doi.org/10.1177/002234336500200104.

Harcup, T., & O'Neill, D. (2016). What Is News? *Journalism Studies*. Advance Online Publication. https://doi.org/10.1080/1461670x.2016.1150193.

O'Reilly. (2010, September 7). *Gov 2.0 Summit 2010: Ellen Miller, "Open Government Scorecard"* [Video file]. https://www.youtube.com/watch?v=UNQteT9Bu2w&t=2m.

Parasie, S., & Dagiral, E. (2013). Data-Driven Journalism and the Public Good: "Computer-Assisted-Reporters" and "Programmer-Journalists" in Chicago. *New Media & Society, 15*(6), 853–871. https://doi.org/10.1177/1461444812463345.

Romano, T. (2017, March 20). In Seattle Art World, Women Run the Show. *The Seattle Times*. http://www.seattletimes.com/entertainment/visual-arts/in-seattle-art-world-women-run-the-show/.

Shenker, J. (2017, July 24). Revealed: The Insidious Creep of Pseudo-Public Space in London. *The Guardian*. https://www.theguardian.com/cities/2017/jul/24/revealed-pseudo-public-space-pops-london-investigation-map.

Sherlock, P., & Bradshaw, P. (2017, July 31). Give MPs Deadline on Hiring Relatives, Campaigners Urge. *BBC News*. http://www.bbc.co.uk/news/uk-england-40709220.

Stoneman, J. (2015). *Does Open Data Need Journalism?* Retrieved from Reuters Institute for the Study of Journalism Website: http://reutersinstitute.politics.ox.ac.uk/sites/default/files/Stoneman%20-%20Does%20Open%20Data%20need%20Journalism.pdf.

Templon, J. (2016, January 17). How BuzzFeed News Used Betting Data to Investigate Match-Fixing in Tennis. *BuzzFeed*. https://www.buzzfeed.com/johntemplon/how-we-used-data-to-investigate-match-fixing-in-tennis.

Usher, N. (2016). *Interactive Journalism: Hackers, Data and Code*. Champaign-Urbana: University of Illinois Press.

Wainwright, D., & Bradshaw, P. (2017, July 18). More Than Two Thirds of Thefts Are 'Never Solved'. *BBC News*. http://www.bbc.co.uk/news/uk-england-40131277.

Williams, A. (2013, May 4). Workshy Map of Britain Revealed: Thousands of Incapacity Benefit Claimants Found to be Capable of Working. *Daily Mail*. http://www.dailymail.co.uk/news/article-2319355/Workshy-map-Britain-revealed-Thousands-incapacity-benefit-claimants-capable-working.html.

Winnett, R., & Rayner, G. (2009). *No Expenses Spared*. London: Bantam Press.

Wu, Y. (2017, March 11). How the Wall Street Journal Visualized the 500+ Conflicts of Interest of The Trumps. *Storybench*. http://www.storybench.org/how-the-wall-street-journal-visualized-the-500-conflicts-of-interest-of-the-trumps/.

Check for updates

CHAPTER 4

Measuring the Unmeasured with Data

Nicolas Kayser-Bril

In 1887, Nellie Bly, an investigative journalist, decided to report on the living conditions in insane asylums in New York City. She went undercover for ten days in an institution at Blackwell's Island (now Roosevelt Island between Manhattan and Brooklyn), feigning insanity herself. There, she observed the personnel, talked to inmates and, upon getting out, wrote a book of her experience, *Ten Days in a Mad-House* (Bly 1887a). The book and the media campaign organised by Bly's employer, the *New York World*, convinced the city of New York to increase funding for asylum wards. In the words of the introduction to the book's edition at the Disability History Museum, it "provided charity commissioners with the ammunition needed to convince city bureaucrats to provide more funding" (Bly 1887a).

One hundred and thirty years later, in 2016, another investigative powerhouse does a story on psych wards in the United States. Buzzfeed's Rosalind Adams showed how a private corporation made a business of locking patients in who had just come for a free mental health assessment, just to bill their insurers (Adams 2016).

On top of numerous interviews, Adams used a series of data sets to show that the situation she was describing did not concern just one

N. Kayser-Bril (✉)
Berlin, Germany

This chapter was originally published on the following website http://blog.nkb.fr/measuring-the-unmeasured under the title Measuring the Unmeasured with Data. It is in the Public Domain.

institution or one region, but the whole of the country, and that the failings were caused by the business practices of a company, in this case, UHS. By contrast, nineteenth-century Bly never once used any statistic to make her case. More strikingly, in the heated exchanges that followed the publication of Bly's investigation, her adversaries did not bring in data either, preferring to focus on details of her stay at the asylum (Bly 1887b).

Why did two investigative journalists act so differently? Why do today's journalists need data to make a case, when a human story was enough a century ago? The nineteenth century did not lack data. The American bureaucracy was already vast in 1887 and Nellie Bly could relatively easily have accessed figures such as the total number of inmates, institutions and doctors who were part of the problem she was exposing. Some investigative journalists of the same period actually used data, when no other investigative technique was available. This is how E.D. Morel researched the slavery system set up by the Belgians in the Congo, for instance, in his book *Red Rubber* (Morel and Johnston 1907).

There are many reasons for this change, but the most important one comes from the new relationship the social elite found to numbers. Starting in the second half of the twentieth century, governments found a new faith in measures. The dreams of a planned society, whether in Soviet countries or in the West, permeated governments and corporations. While government planning has come out of fashion, the search for guidance through numbers remains, bolstered by the advent and promises of the computer and, since the early 2010s, of an artificial intelligence.

This essay explores how journalists and newsrooms use data to measure the unmeasured, thereby remaining relevant in an era of "governance through numbers" (Supiot 2015).

Turning a Story into an Issue

For an investigative project to have as much impact as Nellie Bly's investigation on madhouses, the journalist needs to provide a measure of the issue at play and be able to follow up on this measure through time. Absent such measurements, officials from the bodies responsible for the regulation of the issue might simply argue that the story being described is an anecdote that does not represent the overall situation or that the report is a fabrication. On top of this, they might provide an avalanche

of figures showing how good a job they are doing. Simply providing figures—any figures—enhances the credibility of a report (Tal and Wansink 2016). Had the city of New York known of these communication techniques at the time, it would have easily fought back against Bly's story.

In this environment, some journalists have adapted and became aggregators and curators of databases as well as providers of data-based reports. The processes they follow involve the definition of the issue to be measured through the creation of a methodology, before collecting and structuring data. Such efforts can last several years and have profound impacts.

Together with a consortium of journalists in Europe, we created a database of events during which people had died in their attempt to reach or stay in Europe. The project, called *The Migrants' Files*, ran from 2013 to 2016. Each incident was sourced, located on a map and linked to other information, such as the number of people who died, the number of people who went missing, and the reason for their death. *The Migrants' Files* was not the first project to list fatalities on the borders of Fortress Europe, it built upon previous work by a non-governmental organisation, United Against Racism, that had collected such information since 1993 in an *unstructured* format, unusable for statistical purposes. By structuring the information on the issue, *The Migrants' Files* could provide monthly updates on the number of men and women who died, map the locations of the incidents and perform other analyses, such as computing the mortality rate by route (crossing into Europe from Libya is about a thousand times more dangerous than crossing from Turkey to Greece).

At the date of publication, this approach to the issue of migration was relatively new. Until then, policymakers and journalists had preferred a human interest approach, where the unique tragedy of each mass death event was treated individually (the United Nations High Commissioner for Refugees started collecting structured data on the issue in 2006, but did not release it in any other format than a yearly press release, which might explain why it did not gain traction [United Nations High Commissioner for Refugees 2012]). The data provided by *The Migrants' Files* then served as the foundation for the International Organisation for Migration's (IOM) report on the issue in late 2014, which was followed by a systematic, real-time and sustained effort to provide data on the issue. Once it became clear that IOM would continue its data

gathering operation, *The Migrants' Files* was discontinued. The typical headline moved from "How tragic this shipwreck is" (event-based discourse) to "Death toll in the Mediterranean increases over last year" (the issue *deaths in the Mediterranean* can be measured and reported upon). Without measurement, no one can assess the effect of specific policies on an issue, and therefore no one can hold the powerful accountable for their deeds regarding the issue.

In Spain, the Civio Foundation, a non-profit group of journalists, activists and developers, started the *Indultómetro* in 2012. They collected data from the official gazette on all governmental pardons. While the data was already publicly available and individual cases commented upon, the fact of structuring the data and providing measurements of the issue showed the magnitude of the pardons, as well as how pardons were exchanged for political favours (Bengoa 2015). The long-lasting nature of the project ensured that it gained credibility among other journalists, ensuring coverage, all the while amassing more data that made the possible analyses more comprehensive.

In Germany in 2015, the newsroom of Zeit gathered data on arsons of the homes hosting men and women fleeing war. They showed not only the magnitude of the issue, which so far had been reported as a series of isolated incidents; they also showed that the crimes were almost never investigated by the police (Blickle et al. 2015). While Zeit did not continue updating its data set, another newsroom, taz, took on the effort and provided new data for 2016, following a similar methodology (Stöckel et al. 2017).

While measurements by journalists can create new newsworthy issues, they can also change how existing ones are treated by officials. In the United States, both The Guardian (*The Counted*, 2015–2016) and The Washington Post (*Fatal Force*, 2015–ongoing) created and maintained databases of police casualties. The projects, which recorded about twice the official number of deaths, led to a change in how the Department of Justice monitored the issue (Swaine and McCarthy 2016). The *Iraq Body Count*, which compiles data on violent deaths of civilians in Iraq since 2003, mentions the need to shift to a "human-centred" vision (as opposed to "military-centred" in this case) of the conflict in its rationale, making explicit how data can change how an issue is perceived (Iraq Body Count, n.d.).

Born Out of Necessity

Some of the measurement projects by journalists and citizens are conscious attempts to provide new data on a given issue. *Safecast*, for instance, a project started in 2011 to measure the radiation levels around Fukushima, Japan, was started to provide data free from bias or from "perceived bias" (Ewald 2011). While the project was not launched by a newsroom, its aim of providing the wider public with accurate information to help them reach better decisions aligns perfectly with the goals of journalism, according to more than one definition of the craft.

Not all measurement projects arise from a conscious desire to provide data. *Homicide Watch*, for instance, tracked all homicide in Washington, DC, between 2009 and 2014, and stored the information in a structured format. While the data was ultimately used to provide statistics and shed a new light on the issue, it was launched as a place for people to share information on homicides. The journalist who started the project, Laura Amico, saw that her community had a need for such a place and created it (Amico 2014).

Some projects aiming at providing data on an issue fail to find a use. We ran the project *Rentswatch* between 2015 and 2017, which scraped several thousand classified advertisements for apartments to rent in Europe and provided an average rent price for a given location, precise at the street level. Our hope was that the data could be used to identify the most rapidly gentrifying neighbourhoods and analyse the effects of rent control policies in different countries. The issue we wanted to bring to light was the notion of a *European* housing market, as opposed to 28 national ones, based upon the hunch that many renters live across borders, either as immigrants or cross-border commuters. Despite creating what was probably the largest database of rent prices accessible to journalists, complete with an API (an application programming interface, which helps users access the data easily), the project saw few reusers and no use in investigations (the data was used for entertainment or service purposes). The relative failure of the project shows that a data collection effort on an issue which does not exist as such fails to have any impact if it does not convince fellow journalists of its necessity.

A Counterweight to Official Data

Public institutions tasked with measurements always implement and contribute to a certain vision of the world. Sweden, for instance, is famous for having the longest-running systematic census in the world, which started in 1749. In the eighteenth century, the census collected only the age and social status of the population, because the government was interested in how fast the poorer ranks of society grew. The population was divided into productive and non-productive elements, which included beggars and Sami people. Sami people were then, in 1850, moved from "unproductive" to "non-Swede", together with Roma and Jews. Only in 1993 were the Sami allowed to have a say in how they should be measured. By leaving measurements to the authorities alone, a society implicitly accepts to be ruled by their government's prejudice. Measuring independently is fully part of the idea of keeping power accountable (Kayser-Bril 2016).

The Counted and *Fatal Force*, the databases on people who died of police violence in the United States, and *Safecast* clearly stated their intention to provide another measurement of an issue already measured by officials. In Argentina, newspaper La Nacíon regularly tries to provide alternative measurements of inflation, given that the government's data is untrustworthy. They first aggregated several sources in 2012 (Marshall 2012), then moved on to measuring the number of 500-peso bills in circulation to estimate the increase in the monetary base and derive inflation from there (La Nacíon 2017) (neither attempt was totally successful).

Investigations can also provide a counterweight to official narratives simply by turning a story into a measured issue, as explained above. *Drone Wars*, a project by The Bureau of Investigative Journalism, a British non-profit, was started in 2015 and provides as much structured data as possible on drone strikes carried out by the American and British governments. Analysing the data they gather, the Bureau can offer insights in the wars carried out by these regimes, which sometimes markedly differs from the official discourse. They could show, for instance, that while operations against the Taliban ended officially in 2014, drone strikes against the group continued apace since then (Purkiss 2017).

It happens that journalists simply request access to the official data itself to show that the authorities did not interpret it correctly. An investigation by the German non-profit Correctiv, for instance, showed that

the police of Vienna, Austria, vastly over-reported thefts and drug-related crime in their press releases, while ignoring rape and hate crimes (Kanya and Siebenhofer 2015). Reuters Investigate showed that the United States health authorities did not adequately measure the number of death induced by drug-resistant bacteria (McNeill et al. 2016).

Cost-Effectiveness

Measurement projects rarely require any kind of hardware (exceptions to this are the *Cicada Tracker*, which monitored temperature with custom thermometers, and *Safecast*, which relied on Geiger counters). They usually aggregate information that is available in news reports or in social media and structure it in a database, following a precise methodology. The costs of such projects lie only in the time spent gathering and structuring information, as well as, when needed, the time of setting up a front-facing, interactive interface.

A quick calculation shows that a data collection project done by a journalist costing 75,000€ a year, needing 30 minutes per day, every day ends up costing less than 5000€ a year. Add another 5000€ in yearly costs for a front-end interface and the total yearly cost (notwithstanding the writing of news articles with the data) reaches 10,000€. This estimate fits the order of magnitude of the costs of *The Migrants' Files* (15,000€ over two years) and *Rentswatch* (30,000€ over two years).

Opening data and making it available lets others reuse it and expand the original project's reach. The *Drone Wars* data set, for instance, was used by agency Pitch Interactive in 2013 to create an interactive application, *Out of Sight, Out of Mind*, which was widely shared. Data from *The Migrants' Files* was reused by developer Moriz Büsing to create a visualisation, *15 years*, which was much more powerful than the original one. Considering that many of the newsrooms that create measurement projects are non-profits and that their funders care about "impact", such reuse of data is key to their performance indicators.

ProPublica, a non-profit newsroom which won several Pulitzer prizes, set up an online shop where it sells the data sets derived from their measurement operations. Over the course of two years, they grossed $200,000, or 1% of their operating income (Bilton 2016). Despite this relative success, no other newsrooms followed suit and sold data they collected.

Complementing Academia

The economics of journalistic measurement is fairly insignificant (both the costs and potential revenues are small). More important is its contribution to the production of factual knowledge. Some projects explicitly aim at providing academics with data for them to use. The Hindustan Times, for instance, launched a *Hate Tracker* that monitors instances of hate crimes in India. One of their state goals was to "provide solid, irrefutable evidence to researchers" (*Hindustan Times* 2017). The project was taken offline a few months after it started in a politically-motivated crackdown (Kleis Nielsen 2017), thereby underlining that independent measurements can and does instil fear in political circles.

Even if journalists do not see themselves as competitors to academia, or even as followers of the scientific method, the fact remains that by defining a methodology, making it public and collecting data, journalists engage in science insofar as they create knowledge in a way that can be replicated and falsified.

As academia has proved less able to fulfil its mission of producer of factual knowledge (Kayser-Bril 2017), journalists came to fill the void. Some of the products published by journalists are tailored to look like academic articles, making clear that the level of seriousness the newsroom aims at is equal or greater to the one coming from universities. ProPublica, for instance, published a white paper on the data they collected on the risks of elective surgery which mimics in all ways the style of an academic article (Pierce and Allen 2015).

The findings of journalistic measurement projects can also balance those of academics. ProPublica (again) proved in 2016, using data it curated, that doctors who received payments from a pharmaceutical firm were more likely to prescribe drugs from that firm than those who did not (Ornstein et al. 2016). Two academic studies came to the same conclusions, but they were published after the ProPublica story (Ornstein 2016). Journalists regularly take upon themselves to measure issues systematically, thereby creating exclusive content, reporting events under totally new ways (by making them issues instead of individual stories) and providing a counterweight to official discourse. This not only proves relatively cost-effective (compared to other forms of journalism such as special correspondents), it makes major contributions to the intellectual pursuit of truth, complementing or even surpassing academia in this endeavour. From this laudatory assessment, and taking into account my own biases, one must wonder why newsrooms do not produce more measurements. The answer

flows logically from the previous sentence: very few of them are missioned with finding the truth. It is only logical that non-profit operations such as ProPublica, Civio or The Bureau of Investigative Journalism are at the forefront of journalistic measurements: their mission statements make clear that their job is to find and publish the truth. Other newsrooms have a variety of other missions, from gathering attention in order to sell it to advertisers to providing an oligarch with political influence. For them, it makes no business sense to engage in long-term data collection projects.

References

Adams, R. (2016, December 7). Locked on the Psych Ward. *BuzzFeed*. https://www.buzzfeed.com/rosalindadams/intake?utm_term=.fg832Pa0M#.imkYW8Mwy.

Amico, L. (2014, December 29). Website That Kept Watch on D.C. Homicides Shuts Down. *PBS*. https://www.pbs.org/newshour/show/website-kept-watch-d-c-homicides-shuts.

Bengoa, A. (2015, July 7). La Concesión De Indultos Cae Un 84% Pese a Que Crecen Las Peticiones. *El Pais*. https://politica.elpais.com/politica/2015/07/17/actualidad/1437153240_650862.html.

Bilton, R. (2016, October 7). ProPublica's Data Store, Which Has Pulled in $200K, Is Now Selling Datasets for Other News Orgs. *Nieman Lab*. http://www.niemanlab.org/2016/10/propublicas-data-store-which-has-pulled-in-200k-is-now-selling-datasets-for-other-news-orgs/.

Blickle, P., Biermann, K., Faigle, P., Geisler, A., Hamann, G., Jacobsen, L., … Venohr, S. (2015, December 3). Es brennt in Deutschland. *Zeit Online*. http://www.zeit.de/politik/deutschland/2015-11/rechtsextremismus-fluechtlingsunterkuenfte-gewalt-gegen-fluechtlinge-justiz-taeter-urteile.

Bly, N. (1887a). *Ten Days in a Mad-House*. New York: N.L. Munro. Retrieved from http://www.disabilitymuseum.org/dhm/lib/detail.html?id=1123.

Bly, N. (1887b, October 17). Untruth in Every Line, Nellie Bly Contradicts a Recent Article in the Sun. *New York World*. Retrieved from http://sites.dlib.nyu.edu/undercover/sites/dlib.nyu.edu.undercover/files/documents/uploads/editors/Untruths-in-Every-Line.pdf.

Ewald, D. (2011, March 23). Open Dialogue. *Safecast*. https://blog.safecast.org/2011/03/open-dialogue/.

Hindustan Times. (2017). *About the Hate Tracker*. https://archive.is/20170728074120/hindustantimes.com/hate-tracker/about.

Iraq Body Count. (n.d.). *The Human Cost of War Must Be Recorded*. https://www.iraqbodycount.org/about/rationale/1.

Kanya, E., & Siebenhofer, A. (2015, August 26). Gefühlte Kriminalität. *Correctiv*. https://correctiv.org/recherchen/stories/2015/08/26/polizeiberichte-verzerren-wirklichkeit/.

Kayser-Bril, N. (2016, April 13). *'Free Your Data' Is Over. Now, We Need Data To Be Free* [Blog post]. http://blog.nkb.fr/data-free.

Kayser-Bril, N. (2017, April 10). *The Collapse of Academia* [Blog post]. http://blog.nkb.fr/academia.

Kleis Nielsen, R. (2017, November 15). A creeping quiet in Indian journalism? *Huffington Post Blogs*. https://archive.is/20171212183213/, https://www.huffingtonpost.com/entry/a-creeping-quiet-in-indianjournalism_us_5a046be2e4b0204d0c1714bb#selection-335.0-335.38.

La Nacíon. (2017, March 17). *Tracking the Monetary Base in Pesos 2003–2017, and How This Speaks About Inflation* [Blog post]. http://blogs.lanacion.com.ar/projects/sin-categoria/tracking-the-monetary-base-in-pesos-2003-2017-and-how-this-speaks-about-inflation/.

Marshall, S. (2012, October 25). How Argentina's La Nacion Is Opening Data Without FOI. *Journalism.co.uk*. https://www.journalism.co.uk/news/data-journalism-la-nacion-argentina/s2/a550933/.

McNeill, R., Nelson, D. J., & Abutaleb, Y. (2016, September 7). The Uncounted: The Deadly Epidemic America Is Ignoring. *Reuters*. https://www.reuters.com/investigates/special-report/usa-uncounted-surveillance/.

Morel, E. D., & Johnston, H. H. (1907). *Red Rubber: The Story of the Rubber Slave Trade Which Flourished on the Congo in the Year of Grace 1907*. London: Fisher Unwin.

Ornstein, C. (2016, June 20). Feed Me, Pharma: More Evidence That Industry Meals Are Linked to Costlier Prescribing. *ProPublica*. https://www.propublica.org/article/more-evidence-phrama-meals-are-linked-to-costlier-prescribing.

Ornstein, C., Grochowski Jones, R., & Tigas, M. (2016, March 17). Now There's Proof: Docs Who Get Company Cash Tend to Prescribe More Brand-Name Meds. *ProPublica*. https://www.propublica.org/article/doctors-who-take-company-cash-tend-to-prescribe-more-brand-name-drugs.

Pierce, O., & Allen, M. (2015). Assessing Surgeon-Level Risk of Patient Harm During Elective Surgery for Public Reporting. *ProPublica*. https://static.propublica.org/projects/patient-safety/methodology/surgeon-level-risk-methodology.pdf.

Purkiss, J. (2017, May 25). US Military Operations Against Taliban Dramatically Escalate. *The Bureau of Investigative Journalism*. https://www.thebureauinvestigates.com/stories/2017-05-25/us-military-operations-against-taliban-dramatically-escalate-this-year.

Stöckel, C., Sona, Z., & Bednarczyk, S. (2017, January 2). Es brennt in Deutschland. *taz*. http://www.taz.de/!5367296/.

Supiot, A. (2015). *La gouvernance par les nombres*. Paris: Fayard.

Swaine, J., & McCarthy, C. (2016, December 15). Killings by US Police Logged at Twice the Previous Rate Under New Federal Program. *The Guardian*. https://www.theguardian.com/us-news/2016/dec/15/us-police-killings-department-of-justice-program.

Tal, A., & Wansink, B. (2016). Blinded with Science: Trivial Graphs and Formulas Increase Ad Persuasiveness and Belief in Product Efficacy. *Public Understanding of Science*, *25*(1), 117–125. https://doi.org/10.1177/0963662514549688.

United Nations High Commissioner for Refugees. (2012, January 31). Mediterranean Takes Record as Most Deadly Stretch of Water for Refugees and Migrants in 2011. *UNHCR*. http://www.unhcr.org/news/briefing/2012/1/4f27e01f9/mediterranean-takes-record-deadly-stretch-water-refugees-migrants-2011.html.

Check for updates

CHAPTER 5

Understanding and Contrasting the Data and Exploratory News Outlets

Felipe De La Hoz

Up until relatively recently, the idea of *data journalism* was a compelling yet somewhat mysterious concept in the news industry, a sort of luxury reserved for large newsrooms that could keep around a small stable of dedicated spreadsheet wranglers to perform little-understood data operations and report back their findings in plain English.

Though forms of computer-assisted reporting, or CAR, have been around since the 1960s, recognizable data reporting processes started becoming widespread and professionalized around the 1980s, as the technical tools available to reporters became more sophisticated and accessible. A quintessential example of the early successes in the field is the Pulitzer Prize-winning series on home lending discrimination by Bill Dedman of *The Atlanta Journal-Constitution*. The 1988 effort had Dedman, using techniques pioneered by researchers in academic departments—as well as collaborating with the researchers themselves—obtain data on home loans made over a six-year period in Atlanta and reported to the federal government. Based on this, he conclusively

F. De La Hoz (✉)
New York, NY, USA

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_5

determined that people in middle-income black neighbourhoods were less likely to get loans than counterparts in white neighbourhoods.

That series is very characteristic of the sort of data story of the time: the reporter got their hands on a set of data, in this case through filing a records request with the government; the reporter ran analysis on that data, extracting useful information and identifying patterns; and then the reporter wrote a straightforward news story using that data as one source, bolstering the conclusions drawn from his findings through speaking with other sources.

The picture of data journalism as a field quickly grew more convoluted. As technological advancements continued apace and recognition grew of the wide potential that computational tools unlocked, journalists and publishers began to realize that the classical methods and constraints of news presentation were increasingly less static, and to some degree irrelevant.

The proliferation of available and obtainable data was itself both an opportunity and a conundrum in a society where almost everything had started producing some data footprint, which now lay on top of the existing complexities of government, or business, or sports, or any other beat. Like data journalism, the concept of explanatory journalism had been around for decades, with its tenets first succinctly laid out by The Poynter Institute's Roy Peter Clark in a 1984 essay: write for a general audience, introduce concepts one at a time, use graphics and analogies.

Much of that new framework for crafting coverage sat unused until the sudden information onslaught of the modern information age forced its re-examination as a way to fulfil the core function of helping the society understand itself. It was newly dire to step away from daily, incremental coverage and attempt to present an issue in its relevant context distilled to a summation.

This, in turn, led to the question of the medium and format of the news; after all, the limitations had shifted from what could be printed on paper to what was technically achievable in an internet browser. It was a natural opportunity for data journalism specifically, which seemed uniquely positioned to pioneer new styles. While an interview can only really be presented as text (or videotaped), a series of related facts and figures, such as different sets of numbers per county for the Dedman stories, presents an immense array of possibilities when it comes to packaging. Do you have tabular data? Graphs? Interactive maps?

As part of this general removal of clutter, unnecessary expertise and analysis found itself on the chopping block. There was less need to try to

condense the opinions of an array of ostensibly knowledgeable sources into a meaningful narrative or assertion when the data itself could remove the uncertainty; no point having experts weigh in on whether or not historic patterns of discrimination were on obstacle to black families in Atlanta when you know which side is right.

The advent of modern data and explanatory journalism, then, was a combination of the impulses to condense complicated and information-heavy concepts simply and precisely, present that condensation in formats other than traditional articles—including previously unimaginable formats like interactive maps—and present it authoritatively, with the understanding that certain things are knowable and can be set forth as such.

Different notions of how to best apply these general principles has produced the variety in the outlets that we have come to associate with the genre. This chapter will focus on three of the largest and most fundamental: *The New York Times' The Upshot* vertical; *Vox*, a property of Vox Media; and *FiveThirtyEight*, an ESPN property.

To understand the distinctions between them, it is helpful to examine their inceptions. *FiveThirtyEight* is the oldest of these, having begun as the brainchild of statistician Nate Silver largely as a direct response to his personal dissatisfaction with campaign coverage in the run-up to the 2008 presidential election (Silver 2012). The foundational instinct, in this case, was a strong deference to the straightforward reliability of statistical methods in defiance of mountains of anecdotal, "soft" reporting.

Silver, who pioneered his data-driven style covering baseball, seemed to find no substantial difference between drawing conclusions from batting averages and earned runs over time and doing the same in political scenarios from historical polling and voter behaviour data. The unifying theme to the entire philosophy was that past is precedent, and smart insight is plainly waiting to be gleaned from diligent analysis of the information trail. Silver has not been shy about his disdain for pundits and punditry writ large, with its broad narratives concocted from the liberal interpretation of isolated incidents as opposed to more methodical extrapolation from available data.

Following the 2016 nomination of Donald Trump to be the Republican candidate for president, Silver even wrote a lengthy breakdown of what went wrong with his site's prediction (which had initially given Trump around a 5% chance) titled "How I Acted Like a Pundit And Screwed Up On Donald Trump" (2016). The gist of that piece is that Silver committed his own cardinal sin and presented

predictions without basis in the findings of a specific statistical model. *FiveThirtyEight* has thus remained, among its peers, the most "data-driven" of the outlets, with statistical interpretation continuing to make up its backbone.

It retains and is still largely known for its core concentration on politics (its name is derived from the number of electors in the US Electoral College). After it was started independently by Silver following a stint blogging for the popular Daily Kos in 2008, the site came under the auspices of *The New York Times*, where it was operated as a full-fledged vertical with the title "FiveThirtyEight: Nate Silver's Political Calculus." The site and the entire FiveThirtyEight brand was sold to ESPN in July of 2013 and has since expanded to include sections on politics, sports—a personal interest of Silver's, who began his blogging career at *Baseball Prospectus*—science and health, economics, and culture.

The site has always employed a general format of crisp, straightforward and information-heavy writing supplemented with simple graphics. It has generally eschewed flashy interactive graphical representations in favour of simpler, static visuals that are strict representations of data used to supplement the author's point. Certainly, writing styles differ between its sections, but the thrust throughout is informed analysis and, crucially, speculation.

That brings us to another of the key characteristics of the site: It is foundationally a predictive platform. Its most widely-known accomplishments at the start, after all, were calling elections. The site retains that spirit of forecasting, and many of its articles do feature commentary and prognostication while avoiding the kind of conventional wisdom of the operatives and gurus that Silver so disdains.

The Upshot actually began as a spiritual successor of sorts to *FiveThirtyEight* after Silver and his brand departed for ESPN. It was officially launched in April of 2014 under the direction of David Leonhardt, a long-time journalist and economics writer with the paper who, in a bit of foreshadowing, had launched the analytical sports column "Keeping Score." The top brass knew it couldn't let their blockbuster data vertical decamp to a competitor without replacement, but it was quickly decided that whatever came next could not just be a carbon copy of *FiveThirtyEight* (O'Donovan 2014). There had been grumblings of Silver's clashes with some of the Gray Lady's culture and established political reporting corps (Sullivan 2013); several people at the *Times* had felt that the purely analytical drive was incomplete and did not

incorporate enough "shoe-leather" reporting. *The Upshot* presented the leadership with an opportunity to take the lessons learned and rebuild their data vertical from scratch while trying to create a more equitable marriage between anecdote and data.

One of the most immediately noticeable things about *The Upshot* when compared to its predecessor is that it is attempting to communicate with a more diverse audience in a much more conversational tone. Indeed in his post introducing the site to readers, Leonhardt—who left the section for the opinion page in early 2016 and was replaced by Amanda Cox—wrote that the site would be written "in a direct, plain-spoken way, the same voice we might use when writing an email to a friend" (Leonhardt 2014). The site routinely addresses readers themselves, even presenting some of its articles in the frame of direct advice. One of the site's early hits was an interactive piece titled "Is it Better to Rent or Buy?" by Mike Bostock et al. (2014). The piece presented the reader with a series of interactive sliders that the reader would set to represent their own situation, and the site would automatically crunch the numbers and give the reader a rough recommendation on when they would be better off buying or renting a home.

In this way, *The Upshot* positions itself more firmly in the area of being explanatory and interpretative and is less focused on prediction than *FiveThirtyEight*. Though its pieces tend to incorporate data, not all are purely about data-crunching; the site features analyses of trends, explainers, and criticism, including dissections of other people's data work. Some read like straight news stories about topics that would be of interest to wonks and the technologically inclined, such as a recent datelined piece on how children are being prepared for an automated economy which featured the reporters going to a training program and interviewing parents and experts. This is likely a direct result of the vertical's desire to inhabit the space between pure data and pure shoe-leather.

The rent or buy piece illustrates another defining attribute of the site, which is an openness to experimenting with formats. Oftentimes, graphics are not supplementary but the journalism's main vehicle and they take many forms. Maps are a classic tool in frequent use, but they are joined by four-quadrant graphs, sliders, calendars, and assorted charts; many of them are interactive and responsive to the user. The rent or buy piece is also an example of a style of interactive by *The Upshot* where the reader is able to

enter data, such as by filling out a questionnaire or selecting from a series of options, and has the data crunched and a result spit back out to them.

If *FiveThirtyEight* is fundamentally data-driven and cerebral and natively predictive, *Vox* is the manifestation of a distinct, contrasting journalistic philosophy. It came into being in April 2014 as part of the larger Vox Media organization, which itself was a new company that began as the parent of the sports blog network SB Nation. It is the brainchild of journalist Ezra Klein, who was poached, along with several colleagues, from the *The Washington Post*'s *Wonkblog*, which Klein had founded as a policy and economics vertical within the *Post* with some of the same principles that he would then take to *Vox* (*Wonkblog* continues to operate at the *Post* and itself has its own particular characteristics and operating principles, though we will not delve into it here).

Klein had arrived at the *Post* from a background in blogging; he had developed a degree of wonkish policy expertise before joining the newspaper in 2009, and was given wide latitude to build *Wonkblog* up into a data-heavy traffic driver and must-read in D.C. political circles, with a roughly similar degree of autonomy as Silver had at *FiveThirtyEight*. He was lured away by Vox Media after failing to convince the leadership at the *Post* to provide him with a revamped platform that would have included an expanded staff (Byers and Gold 2014); after the move, Klein stated one of the catalysts was technology and the lack of a content management system that could allow robust storytelling in the way he was envisioning (Kaufman 2014).

It does not take a seasoned media analyst to identify the Vox's central intent. After all, it's right there in its tagline: Understand the News. It takes to heart the tenets of Clark's Poynter essay and essentially attempts a distillation of current events. *FiveThirtyEight* acts as a microscope of granular data crunching and prognostication, while *Vox* prefers to be a wide-angle lens trying to snapshot the whole picture of each consequential issue it tackles. It is no less an essentially data-focused enterprise but is not as preoccupied with the statistical dissection of each new development in an important story as it is with placing it in context. Perhaps the most concrete and celebrated manifestation of this are the site's *card stacks*, which are a kind of living interactive article that expands as stories develop; instead of a series of static articles or updates, the stack is more akin to a marriage between Wikipedia and a traditional piece of journalism. It will not just tell you that something happened and that it matters, but it aims to tell you why it matters, what historically set the stage for

it to matter, and how it might continue to matter. A notable example of this is the stack "Everything you need to know about Israel-Palestine" (Beauchamp 2017). The topic is certainly among the most multifaceted, impenetrable conflicts in modern history, with unending layers of causes and effects rooted in overlapping historical, religious, cultural, and martial contexts. A casual reader arriving at any isolated story on the Israeli–Palestinian conflict can expect to be hopelessly lost in the effort to grasp the development's significance for the larger state of affairs, and this in itself can be a powerful disincentive to dig deeper; why bother, it is practically Gordian.

So *Vox* truly starts at the beginning: the introductory card is "What are Israel and Palestine? Why are they fighting?" and the second card tackles "What is Zionism?" Even someone with no knowledge base on the subject can read and follow that. From there, each subsequent card builds on itself, arriving at topics like "Why are the US and Israel so friendly?" (card 13) and eventually getting to specifics such as "Why did Israel and Hamas go to war in July 2014?" (card 19). Ideally, a person could arrive with very little comprehension and leave having derived a decent understanding of the conflict's trajectory, both past and potentially where it could be headed in the future. By necessity, these overviews do not go very much in depth; each card tends to hover at around 400 words in length. This approach is not meant to supplant traditional news stories on ongoing developments on the topic, but rather to act as a constantly updating appetizer that will allow people to properly digest the news as it comes.

The site is also a breeding ground for experimentation in news presentation. Its explanatory essence leads it to use everything from standard graphs to cartoons in making news digestible for the unenlightened reader. Whatever the editors believe to be the most straightforward way to convey information is the medium they'll use. The card stacks are one example, but they are far from the most unconventional of *Vox*'s news presentations.

To get a more concrete sense of how the operational principles of each outlet shape their approach, it is useful to consider their handling of the same story. For our purposes, an ideal candidate is the 2017 debate over the repeal of the Affordable Care Act, specifically the Senate push to pass the so-called Graham–Cassidy bill to gut Obamacare. It was a prominent national story on a very complex topic involving lots of data and moving parts and intersected with health care, politics, taxes, and the

legislative heft of a federal administration newly controlled in its entirety by the Republican party. In short, the sort of thing that lends itself to plenty of deep, technical, and explanatory reporting.

The first *FiveThirtyEight* story to mention Graham–Cassidy was titled "Republicans Really Could Repeal Obamacare This Time," and written by senior political writer Perry Bacon, Jr. (2017). The piece is classic *FiveThirtyEight*, starting with the well-informed prognostication its title lays out. It is packaged as modules of information, such as section of brief bullet-pointed highlights of the bill's specific intended consequences; a quick explanation of why the bill is being introduced after previous efforts failed and why Republicans had to pass it before a September 30 deadline; and a detailed section on the likelihood of different Republican senators voting in favour, structured as four segments (almost certain yes, likely yes, almost certain no, likely no, and wild cards). It is number-heavy and detailed in a direct way, offering up analysis without wasting many words. For maximum effect, the reader would have to have some prior knowledge of the circumstances surrounding the legislative push; parachuting in would not quite work.

Another piece, "The New Health Care Bill Is Not So Great For States That Like Obamacare" by Anna Maria Barry-Jester (2017), was focused on taking a popular argument in favour the bill—that it would be uniformly beneficial for individual states in granting greater flexibility—and systematically blowing holes in it. It crunches the numbers on Medicaid funding changes and matches the information with states that expanded the Medicaid program under Obamacare, all presented using one of the site's characteristic multichromatic graphics. The piece finishes by pointing out that states with successful insurance exchanges will have to build up a new Medicare-replacement infrastructure with less money than they had before. In sum, it pairs data analysis with the institutional know-how to plainly assert what will happen, above the fray of politically motivated declarations of what *could* happen and inserts some predictions on the policy aftereffects of the bill's passing beyond the obvious.

True to form, coverage in the *Upshot* leans more towards the explanatory rather than a pure analytical system. Included in this is some of its hallmark pragmatist and utilitarian approach, such as discussing how the demise of the bill and the larger repeal effort could directly affect members of the public with a piece titled "How Failure of the Obamacare Repeal Affects Consumers," by Margot Sanger-Katz (2017b). The piece touches on but largely sets aside the political manoeuvring and the

big-picture policy implications and focuses on what the bill's death will mean in practice for different categories of people. For example, it reads "Those with employer insurance can also coast along with the status quo; there have been no policy changes affecting the price or comprehensiveness of workplace health plans."

As the legislative fight was ongoing, the vertical also attempted to meticulously dissect what would occur if Graham–Cassidy were to be passed, in "How the Latest Obamacare Repeal Plan Would Work," by Reed Abelson and Sanger-Katz (2017); in contrast to *FiveThirtyEight*'s method, it is more conversational and seems geared towards the layman. "The legislation also eliminates the tax penalty that people who refuse coverage face — the individual mandate — which could discourage insurance enrolment among healthier people, who are critical to making the program work," the authors explain. They go through some potential impacts of the block grant system, the spectre of a consumer protections rollback, and ramifications for Medicaid. It's designed to get the reader up to speed on what to expect if the bill becomes law, without diving too deep into each aspect.

The *Upshot* also cleverly exposed the political dynamics at play above and beyond the bill's technicalities by highlighting how many of the concerns raised by senators pushing the bill could have been addressed through provisions already in the ACA itself in a piece titled "One Way for G.O.P. to Achieve Some Repeal Goals? It's Already Part of Obamacare," by Sanger-Katz (2017a). In doing so, it further leads the reader to an understanding of the environment in which the debate is taking place, one in which not only outcomes matter but sometimes process and political points matter even more.

Vox took a slightly different approach; the site conducted saturation coverage of the Graham–Cassidy fight, with the sort of incrementalist approach that it was founded on. This includes at least three separate headlines containing the word "explained"—"Graham-Cassidy: The last GOP health plan left standing, explained," by Sarah Kliff, "The past 24 hours of Graham-Cassidy, explained," by Dylan Scott, and "Graham-Cassidy, the GOP's last health care proposal, explained with a cartoon," by Alvin Chang (2017). As its title suggests, that last one is formatted as a series of cartoons with accompanying explainer text. Chang uses the example of feeding birds as a representation of the different insurance systems, and keeps the cartoons and concurrent reporting simple and to

the point. It is meant for an audience that quite possibly had no idea how Obamacare or the Graham–Cassidy bill really worked.

The site also takes a separate tack in the tenor of the coverage, which often takes on a chattier, more first-person tone. "I've covered the GOP repeal plans since day one. Graham-Cassidy is the most radical" reads the headline of a Sarah Kliff piece (2017), which uses the author's impressions based on her own reporting trajectory combined with multiple data sources and experts to conceptualize for the reader what the likely impact of the bill's passage would be. Kliff also makes more analytical, subjective declarations not based strictly on numeric data, such as the assertion that the bill is moving along despite actually making issues in the previous bills worse, largely on the basis of it being a last-ditch effort. Another piece, "GOP Senators are rushing to pass Graham-Cassidy. We asked 9 to explain what it does" by Jeff Stein (2017), seems to have a bit of a tongue-in-cheek premise at the surface, but in reality, it provides the valuable frame of reference that legislators pushing the bill don't necessarily agree on or even know what it does. It is the sort of idea that seems too obvious to do, yet yields valuable background to the ongoing debate.

References

Abelson, R., & Sanger-Katz, M. (2017, September 20). How the Latest Obamacare Repeal Plan Would Work. *The New York Times*. Retrieved from http://www.nytimes.com.

Bacon, P., Jr. (2017, September 18). Republicans Really Could Repeal Obamacare This Time. *FiveThirtyEight*. Retrieved from http://www.fivethirtyeight.com.

Barry-Jester, A. M. (2017, September 20). The New Health Care Bill Is Not So Great For States That Like Obamacare. *FiveThirtyEight*. Retrieved from http://www.fivethirtyeight.com.

Beauchamp, Z. (updated 2017, May 1). Everything You Need To Know About Israel–Palestine. *Vox*. Retrieved from http://www.vox.com.

Bostock, M., Carter, S., & Tse, A. (2014). Is It Better to Rent or Buy? *The New York Times*. Retrieved from http://www.nytimes.com.

Byers, D., & Gold, H. (2014, January 21). Why the Post Passed on Ezra Klein. *Politico*. Retrieved from http://www.politico.com.

Chang, A. (2017, September 20). Graham-Cassidy, The GOP's Last Health Care Proposal, Explained with a Cartoon. *Vox*. Retrieved from http://www.vox.com.

Clark, R. P. (1984, January–February). Making Hard Facts Easy Reading: 14 Steps to Clarity. *Washington Journalism Review*. Retrieved from http://cdn.nycitynewsservice.com/blogs.dir/422/files/2014/02/MakingHardFactsEasyReading.pdf.

Dedman, B. (1988, May 1). Atlanta Blacks Losing in Home Loans Scramble; Banks Favor White Areas by 5-1 Margin. *The Atlanta Journal-Constitution*. Retrieved from http://www.powerreporting.com.

Kaufman, L. (2014, April 6). Vox Takes Melding of Journalism and Technology to a New Level. *The New York Times*. Retrieved from http://www.nytimes.com.

Kliff, S. (2017, September 20). I've Covered the GOP Repeal Plans Since Day One. Graham-Cassidy is the Most Radical. *Vox*. Retrieved from http://www.vox.com.

Leonhardt, D. (2014, April 22). Navigate News With The Upshot. *The New York Times*. Retrieved from http://www.nytimes.com.

O'Donovan, C. (2014, April 23). Q&A: David Leonhardt Says The Upshot Won't Replace Nate Silver at *The New York Times*. *Nieman Lab*. Retrieved from http://www.niemanlab.org.

Sanger-Katz, M. (2017, September 23). One Way for G.O.P. to Achieve Some Repeal Goals? It's Already Part of Obamacare. *The New York Times*. Retrieved from http://www.nytimes.com.

Sanger-Katz, M. (2017, September 26). How Failure of the Obamacare Repeal Affects Consumers. *The New York Times*. Retrieved from http://www.nytimes.com.

Silver, N. (2012, November 10). Why I Started FiveThirtyEight. *The Daily Beast*. Retrieved from http://www.thedailybeast.com.

Silver, N. (2016, May 18). How I Acted Like a Pundit and Screwed Up on Donald Trump. *FiveThirtyEight*. Retrieved from http://www.fivethirtyeight.com.

Stein, J. (2017, September 20). GOP Senators are Rushing To Pass Graham-Cassidy. We Asked 9 to Explain What It Does. *Vox*. Retrieved from http://www.vox.com.

Sullivan, M. (2013, July 22). Nate Silver Went Against the Grain for Some at The Times. *The New York Times*. Retrieved from http://www.nytimes.com.

Check for updates

CHAPTER 6

Uncovering International Stories with Data and Collaboration

Emilia Díaz-Struck and Mar Cabra

1997 was a year where politics dominated much of the news. Bill Clinton started his second term as President of the United States in January. In May, Tony Blair became the first Labour Party Prime Minister in 18 years. Diana, Princess of Wales, died in a car accident in Paris in August.

In October that year, the International Consortium of Investigative Journalists was born as a network of journalists who collaborated on in-depth investigative stories and as a project of the nonprofit Center for Public Integrity. Unlike any of the other news events, this act did not make the front pages of the newspapers, but it created a ripple effect that would impact politics and policy around the world almost twenty years later in an unprecedented manner. It would also transform journalism, proving the enormous power of cross-border collaborations.

E. Díaz-Struck (✉)
The International Consortium of Investigative Journalists (ICIJ), Washington, DC, USA

M. Cabra
The International Consortium of Investigative Journalists (ICIJ), Madrid, Spain

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_6

In April 3, 2016, ICIJ broke news with the Panama Papers, an investigation into the offshore holdings of world political leaders and details of the hidden financial dealings of fraudsters, drug traffickers, billionaires, celebrities, sports stars and more. It was the largest collaboration in journalism history at the time, involving more than 370 reporters in nearly 80 countries. Together, they sifted through a leak of 11.5 million files originally obtained by German newspaper *Süddeutsche Zeitung* and then reported jointly and separately in secret for a year.

The more than 100 media organisations involved had published 4700 articles by December 2016, which led to historic global events. There were resignations by high-ranking officials, including the prime minister of Iceland and the prime minister of Pakistan. Eight months after the investigation was published, ICIJ accounted for at least 150 inquiries, audits or investigations in 79 countries due to its revelations. ICIJ won almost 20 awards, including the Pulitzer Prize and the Data Journalism Award.

A year and a half later, in November 2017, the ICIJ and its media partners dominated the headlines again with another project on the offshore economy. The Paradise Papers consolidated the organisation's radical sharing model, involving more than 380 journalists on six continents. It revealed the offshore interests and activities of more than 120 politicians and world leaders, including Queen Elizabeth II, and 13 advisers, major donors and members of President Donald J. Trump's administration. It also exposed the tax engineering and avoidance of more than 100 multinational corporations, including Apple, Nike and Botox-maker Allergan.

The Panama and Paradise Papers' success recipe was not created overnight. The methods and techniques used to achieve such impact through journalistic collaboration are a result of trial and error, project after project, since 1997. Over this time, one of the main aspects ICIJ worked on was building trust within its network of reporters.[1]

Collaborations do not work without trust. ICIJ's model requires reporters to share leads and findings, which multiplies the power of the investigative team across borders.

Sharing is not the natural step for many journalists. In the media, bylines and scoops often have been valued too highly to share the glory with others. Collaborating is also extra work, so why go through the

[1]The full team that worked on the Panama Papers is here: https://panamapapers.icij.org/about.html and the Paradise Papers team here: https://www.icij.org/investigations/paradise-papers/about/.

hustle? Because the world is turning more and more complex and the issues we investigate tend to connect to several countries. Sharing is the only way to deal with this new reality, which also involves sorting through massive amounts of information and document dumps.

Corporations and organised crime work on a global scale. ICIJ showed that journalists can too. Its recipe for cross-border collaborations has 10 steps. When followed correctly, the results can be long-lasting and world-changing as in the investigations of the Panama and the Paradise Papers. We encourage you to try it.

Have a Clear Project

Before jumping into an international collaboration that involves journalists around the world, the first thing is to choose the right project. ICIJ takes this phase seriously, as working on investigations takes time, and it is important to have clarity about the project before reaching out to partners.

There are three key questions that ICIJ takes into account to decide which topic to investigate: Is it a global issue? Are people affected by it in a systematic way? Is it possible to create change?

By asking these questions, which are aligned with the organisation's mission, ICIJ determines whether the possible investigation is in the public interest and evaluates its viability.

The initial research could take months. Once the topic has been selected, ICIJ starts inviting partners to work together on the project.

A role that is key in this process—and during the whole course of the investigation—is that of the project manager. ICIJ deputy director Marina Walker was the project manager of both the Panama and the Paradise Papers investigations, joined by director Gerard Ryle in the former project. It is important to have a person that makes sure everyone in the team collaborates, following a set of ground rules that are agreed to at the beginning.

Develop a Collaborative State of Mind

People are one of the most important elements of ICIJ projects—and any collaboration, so choosing your partners is another key step.

Bringing the right journalists on board is like organising a dinner party, as Marina Walker likes to say. To avoid ruining the party, the combination of all the individuals has to work, and there should be trust among everyone involved. ICIJ first searches for partners inside its network

of members, who are all respected investigative journalists in their own countries. It also reaches out to other investigative journalists who have approached the organisation and shown interest in participating in future projects. Sometimes, media organisations are approached directly.

In the process of choosing the right partners, ICIJ looks for journalists who like to share. ICIJ's model is based on trust and radical sharing. If one journalist gives out information but does not receive anything back from other team members, trust is easily lost. ICIJ also searches for regional diversity and involves small and big media partners. What all of them have in common are their great investigative skills and high journalism standards. They are also happy to share findings and documents with colleagues from different parts of the world and bring their own expertise to the global collaboration.

The Panama Papers was kept a secret by all the team members for a year, up until publication. Everyone understood the importance of the work and the team acted like a family. Revealing the information about the project ahead of time could affect everyone on the team, as well as jeopardise the whole investigation.

To facilitate the development of a collaborative state of mind, there should be a set of ground rules on how the collaboration will work that everyone agrees to when joining a project. While working together on the investigation, common goals and deadlines are established, as is the publication date.

Each global project brings new challenges. ICIJ innovates while collaborating and working on specific investigations, searching for answers for the project's needs as well as partners' needs. Through data and technology, the organisation tries to find ways of using platforms that will facilitate investigations into complex topics like corruption that have global connections.

Work Across Borders

Every project ICIJ does involves journalists from several continents, which means most are not native English speakers and come from different cultures. While the project manager oversees the general execution of the investigation and coordinates the collaboration, ICIJ has found through the years that having regional coordinators helps mitigate the cultural differences and better assist with local needs. For the Paradise Papers, ICIJ had regional coordinators who covered North America, Latin America, Africa and the Middle East, Europe, Asia and Oceania.

Their main task is to keep in constant communication with journalists from their regions and to follow-up with them on the progress of their reporting. One important aspect of this coordination is to help identify connections between different countries that could lead to more collaboration within that region and with other regions.

There are also other ways in which trust and communications are reinforced. ICIJ does at least one in-person meeting per project to talk about the investigation, how the team is going to work together, and about initial leads that would require more digging and reporting, among other things. The Panama and the Paradise Papers were each one-year projects. In both cases, the main meeting was held at the headquarters of *Süddeutsche Zeitung* in Munich, Germany, around seven months before publication. In addition to that in-person meeting, online meetings and training also take place.

One of ICIJ's innovations for cross-border collaborations has been to use technology to strengthen communication. The organisation's data team created the *Global I-Hub*, a private website that works as a social network to allow journalists to share information and leads, securely.

In this virtual newsroom, reporters can create groups around issues of interest so they can collaboratively organise reporting on specific topics. Journalists also share tips to improve research. In the Panama and Paradise Papers, for example, there was a group called "The Search Cooperative." Through the *Global I-Hub*, team members can also share files and exchange private messages with other collaborators. In sum, this platform provides a one-stop shop where reporters around the world can be in touch all the time and know what others are working on, regardless of how far they are from each other. This way, they can identify when a specific story is of interest to other countries.

Add Data

The importance of data in investigations has been growing exponentially in the past years. On one side, leaks have become bigger and bigger. On the other side, data has become more and more complex.

The Panama Papers, for instance, involved 11.5 million files and the Paradise Papers 13.4 million. The leaks that originated the projects contained in both cases a large number of documents with complicated financial information on companies registered in tax havens. The data revealed on many occasions the true ownership behind corporate

structures, which would have been impossible to gather using public data. It also involved convoluted operations aimed at reducing taxes for big corporations.

ICIJ needed to incorporate in its team new expertise to handle these large sets of data and make them available in a secure way to its network of partners. It was just in the spring of 2014 that it created its Data & Research Unit, a multidisciplinary team of data journalists, database engineers and developers, to work on developing technologies to facilitate the sharing and exploration of documents and to perform data analysis for ICIJ's projects.

The increased size of leaks and the relevance of the data inside them highlight the importance of the relationship with and the protection of whistleblowers. The creation of secure platforms to share leaked information and the use of encryption helps keep the sources safe. ICIJ had to train reporters around the world on cybersecurity techniques for its last projects.

However, leaks are not the only way for journalists to find data that can be useful for stories. Public data also plays an important role in investigative journalism, and it can also come in large sizes. For example, in the Panama and the Paradise Papers public corporate registries were—when available—important to verify some of the information in the leak, such as if certain companies existed in those jurisdictions or whether the incorporation date matched between public and leaked data. In addition, other public sources such as court records and property records played a vital role in the research and reporting done in connection with the leaked documents.

Public data can also trigger new projects. That was the case of ICIJ's 2015 investigation *Evicted and Abandoned*, where through public reports from the World Bank, ICIJ was able to account for more than 3 million people either physically or economically displaced by nearly 1000 World Bank-financed projects between 2004 and 2013.

Spend Time Looking and Understanding the Data

The research, reporting and data analysis that takes place in journalistic investigations with a strong data component needs a thorough approach to the data.

It is key not to take things for granted. When the data is reviewed, it has to be explored with a critical eye. The process—beyond making it

searchable and accessible to the network of collaborators—involves, on the one hand, verifying the authenticity of the information and on the other, understanding the data.

Rushing things is not the best approach! Complex data sets require time to check and better understand what kind of information is available inside them, as well as the best ways to explore it.

In the case of big document dumps, it is important to check first what type of files are included to explore the best ways to organise the data inside them, how to make the files available to everyone and whether the information from certain documents can be structured.

Are there ways in which technology can help to better understand and explore the data? That is a question that remains open throughout the investigative process and which ICIJ keeps constantly in mind during projects, using innovation to facilitate the collaboration between partners.

Once the files have been shared with the team, it is also important to explore the documents and the data thoroughly to identify aspects that could be relevant to the research. In the Panama and Paradise Papers, we used specific terms that are part of the offshore lingo to make searches more effective and identified types of documents that contained information that could lead to useful findings—such as financial statements, board meetings, and shareholder certificates. ICIJ shared the results of this exploration with the network, which in turn shared ways the journalists had found to explore the documents that turned up strong leads.

Last but not least, all the data work should be combined with traditional investigative reporting skills. Conducting interviews, finding good expert sources and going to the field to see how the data connects with the world are also key to understanding what the full story is. A combination of shoe-leather reporting, data reporting and analysis, and the work of developers facilitates the production of important public interest journalism through clues that are hidden in millions of files or rows of a database.

Analyse the Threats to Your Investigation

The world changed in June 2013, when whistleblower Edward Snowden revealed in *The Guardian* and *The Washington Post* the global surveillance programmes of the United States National Security Agency. Before then, only a small group of journalists were expert in cybersecurity.

Now, knowing how to use encryption and basic security skills are a must for many investigations. This new reality also means that at the beginning of every investigation—especially if it is a collaboration—one of the most important questions to answer is who is a threat to the project.

Threat modelling is a standard process followed in information technology to identify potential threats and vulnerabilities in a piece of software or a project. The technique can easily be applied to journalism. In its *Surveillance Self-Defense* guide, the Electronic Frontier Foundation identifies five questions to answer (Electronic Frontier Foundation 2017): what do I want to protect?; who do I want to protect it from?; how bad are the consequences if I fail?; how likely is it that I will need to protect it?; and how much trouble am I willing to go through to try to prevent potential consequences?

In ICIJ, we found that following strict security protocols for every single case is not the best approach, as the more security you bring to a project, the more complex it is technologically. In order to carry out collaborations like the Panama and the Paradise Papers, we had to make compromises. For example, we determined that government agencies would not be our biggest enemies in the case of potential surveillance, and this allowed us to use the internet to store and distribute our data to partners.

At the same time, we took some scrupulous measures, such as forcing two-factor authentication in all of our platforms, requiring journalists to enter a temporary code from their phone when they logged in as well as their passwords. Security steps like this, or forcing encryption in project communications, required ICIJ to train team members. But time spent on this was well spent because no matter how many security steps one takes, people are always the weakest link in the chain.

Share, Radically

ICIJ believes the only way to collaborate is through radical sharing. Everyone in the team has access to the key project materials and documents if the project is based on a leak. It also means partners are expected to share leads and their research—some even share transcripts of interviews. Everything counts towards crafting the best and most accurate story possible.

All the materials ICIJ produces for publication as a newsroom—articles, graphics, art, etc.—are provided to the media participating in the

project ahead of publication, so they can use them however they prefer, translating or adapting them to their own audience. Other partners do the same. For example, those working in television contribute raw footage for others to use. The one rule is that anyone who publishes the content is the one legally responsible for it.

Besides the articles, the other most popular items that ICIJ creates are numerical breakdowns and multimedia products. The Data & Research Unit gathers those figures that speak globally about the investigation, such as the number or type of files, the countries involved or the size of the leak. Some are more project-specific, such as the number of banks that created offshore entities. Radical sharing also leads to staggering figures. Thanks to the leads shared by the global team of reporters in both the Panama and the Paradise Papers, we accounted for more than 250 politicians and world leaders connected to companies in tax havens.

All the departments of the ICIJ are involved from the beginning of the project to come up with the best ways to tell the story—text is just one of them. In the Panama Papers, for example, we did an animated video to explain who the victims of the offshore system are (ICIJ 2016a), which has been watched by almost two million people. Interactive graphics helped us explain complex topics in a simpler way. With *The Power Players* (ICIJ, n.d.-a), we told the stories of the politicians mentioned in both leaks. *Stairway to Tax Haven* (ICIJ 2016b) was a game to understand how the offshore world works, developed in collaboration with French newspaper Le Monde. *The Influencers* reveals step by step how Trump's inner circle is connected to tax havens. All these products are shared with partners and translated into dozens of languages.

Use Technology to Boost Journalism

The *Global I-Hub*, ICIJ's virtual newsroom, allows reporters to be constantly in touch with each other throughout a project, in the same way Facebook does for communities of friends and family. But it is not the only technology ICIJ uses to disrupt old investigative reporting models and facilitate collaboration across borders.

In both the Panama and the Paradise Papers, ICIJ adapted existing open-source technology to provide a secure platform where team members could investigate the leaked documents in full. The files were processed and indexed so that journalists could make searches in a similar

way to searching in Google. Research was also accelerated with the possibility of uploading lists of names to see if there were any hits.

Visualisations can also help reporting. We partnered with tech companies *Neo4j* and *Linkurious* to obtain software that allowed our reporters to easily understand who was connected to whom inside the leaks.

In the future, new needs will emerge and ICIJ will probably apply other technologies to meet them. Regardless of the software, two things will remain true: New technologies require training team members on them, and managing access takes time. ICIJ had to create a "support team" just to help journalists with problems over our platforms. We also created manuals and conducted training in three languages for people on four continents.

Empower Readers

Collaborations do not need to end at publication. In ICIJ we learned that by sharing large quantities of the project's data with the public, we gave a second life to the investigation. For privacy and source protection, we could not dump all the leaked files in the internet. We ended up making searchable in a website the names of the people connected to companies and trusts in tax havens.

The Offshore Leaks database (ICIJ, n.d.-b) received 12 million visits after the Panama and Paradise Papers data was incorporated in it, and it is used regularly by academics, NGOs and tax agencies. By connecting their data to ours, governments are recovering millions of dollars in unpaid taxes. Spain retrieved USD$122 million and South Korea USD$120 million, just to mention a couple of examples.

Authorities are also discovering new leads for investigations. Europol, Europe's law enforcement agency, matched our offshore database with its own and found more than 3000 probable matches to organised crime and tax fraud. Of those matches, 116 were related to Islamic terrorism.

Journalists who were not part of the initial investigation are also finding new stories in the publicly available data and approaching ICIJ to start new collaborations. This constant discovery may never end because as the world moves on, new people become public figures, unexpected politicians get into office, or we learn about the names of criminals that operate in the shadows, meaning more and more connections to investigate.

Be Prepared for New Challenges

If there is anything certain at ICIJ after 20 years of cross-border collaborations, it is that no two investigations are the same. Even the Panama and the Paradise Papers, which may seem similar in procedures and content, posed different challenges. From the data perspective, for instance, the Paradise Papers was more knotty, involving file formats such as PSTs (compressed email files), which were difficult to unpack and index. The documents were also more complicated to understand, as they included convoluted tax schemes that only experts could untangle well.

In between both investigations, ICIJ had a challenge of its own. In February 2017, it spun off from the Center for Public Integrity, becoming an independent nonprofit news organisation. "We believe this new structure will allow us to extend our global reach and impact even farther," said Gerard Ryle when the announcement was made.

At the end of 2017, the future seems bright for journalistic collaborations. There are more projects being done collaboratively, new organisations that promote sharing among reporters are being formed and the topic is even being studied at the university level. Some of the challenges ahead are how all these new networks will interact with each other, and also how journalists will start sharing documents on a more regular basis, not just while working on specific projects. Time will tell. We will have to write another book chapter in a few years explaining how we solved them.

References

Electronic Frontier Foundation. (2017, September 7). *Assessing Your Risks*. Retrieved from https://ssd.eff.org/en/module/assessing-your-risks.

ICIJ. (n.d.-a). *Explore Offshore Connections of Political Power Players*. Retrieved from https://offshoreleaks.icij.org/stories.

ICIJ. (n.d.-b). *Offshore Leaks Database*. Retrieved from https://offshoreleaks.icij.org/.

ICIJ. (2016a, April 3). *The Panama Papers: Victims of Offshore* [Video file]. Retrieved from https://www.youtube.com/watch?v=F6XnH_OnpO0.

ICIJ. (2016b). *Interactive Game: Stairway to Tax Haven*. Retrieved from https://www.icij.org/investigations/panama-papers/stairway-tax-heaven/.



CHAPTER 7

Data-Driven Human Rights Investigations

Alice Kohli

In this chapter, I would like to explore to what extent the developments of digital investigative journalism have reached human rights fact-finding—and where the new techniques fall short in face of the reality of an unstable political environment.

What Is Human Rights Fact-Finding?

A multitude of people and organisations are involved in research and fact-finding missions in areas where human rights violations are suspected. Investigations are conducted by non-governmental organisations, but also by the judicial systems or the ombudsman of the respective states, as well as by supranational bodies such as the United Nations Organisation. For example, the work of UN rapporteurs or truth commissions assigned by International Criminal Courts to investigate war crimes during or after armed conflicts fall under this category.

A. Kohli (✉)
Public Eye, Zürich, Switzerland

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_7

What Is Relevant for Journalists?

Truth commissions that work in a legal setup are not always willing to disclose their information to journalists—fearing it might contest their trial—as they hope that it will lead to the conviction of a culprit before a court. So, in some circumstances, it can even be harmful to their cause to involve the news media. On the other hand, sources provided by the government can pose other problems for journalists: in a state of crisis or war the government or administration might not be fit for the task of investigating human rights violations on its own, or worse—be itself directly involved in the abuses. This is why media often rely on the work previously conducted by non-governmental organisations (NGOs) or advocacy groups when researching human rights violations. For this reason I will focus on the work conducted by NGOs in this chapter.

NGOs and Journalism

Amnesty International was the first NGO to pioneer fact-finding in the early 1960s, according to British human rights barrister Geoffrey Robertson (2010, p. 21). Since then, NGO fact-finding has often overlapped with journalistic work and entailed similar techniques. Organisations such as Amnesty International or Human Rights Watch have made it their goal to shed light on the often pitiful situation of basic human rights in various political and geographical contexts and to put pressure on policymakers around the globe to act. Oftentimes, NGOs do not only seek a journalistic outlet to publish their research—they also investigate side-by-side with reporters directly affiliated to news media in order to gain momentum for promoting their political demands.

Compared to journalists, NGO researchers usually have a long lasting network in their field of expertise. Depending on their work, they are either rooted within the local community or in an international policy setup. They have a deeper insight into the multiple complex causes of a problem than journalists normally do. Additionally, human rights organisations often have more time and funds to investigate the issues at hand.

NGOs and Data Journalism

In the last decades, some organisations have started to use data analysis techniques to research human rights violations. A few have even specialised in using information technology to meet the need of quantitative

analysis in human rights investigations. Nevertheless, the buzzword "data journalism" or "data-driven investigation" has not fully caught on in the world of NGOs. Asking around at my own workplace—the Swiss human rights NGO Public Eye (formerly known as Berne Declaration)—proves that the term "data journalism" does not cause a spark of associations among human rights fact-finders. Many of them have worked at the organisation for years and have—understandably—only loosely followed the discussions in media journalism.

Even if they are rarely referred to as a special technique, data investigations are conducted at Public Eye as a matter of routine, just as expert interviews and on-the-ground research are. One of the last major investigations conducted by Public Eye resulted in a report titled "Dirty Diesel" (Public Eye 2016) on an unethical business model of Swiss commodity traders in Africa. The companies mentioned in the report deliberately lowered the qualities of diesel and gasoline to take advantage of the weak fuel standards they encountered on the continent—ironically, where they had sourced the commodity in the first place. In consequence, people in West African countries have been suffering from severe health damages, such as respiratory diseases. During the research leading to the report, data journalism techniques were used in various circumstances.

"We used international trade databases, such as UN Comtrade or national import and export statistics to gain insight into the fuel business", recalls Gian-Valentino Viredaz, researcher at Public Eye. Through thorough database research, his team established the sulphur content of the diesel fuel exported as well as the destined recipients. It was partly on that basis that they identified the priority countries for on-the-ground investigations. Marietta Harjono, an external researcher who worked on the project, skillfully used ship tracking and chartering software to find out about the freight on the vessels entering and leaving the Netherlands. For her, the database investigations were useful supporting tools. "But for this kind of topic, which has never been investigated before, you have to go out and leave your desk."

To be able to understand the intransparent business model of intentional blending of low quality or unhealthy blendstocks into gasoline and diesel required the team to be very creative and use different techniques to get the bigger picture. "To go into the field, find people willing to speak, talk to insiders, seamen, blending and refinery experts, petrolab supervisors, ship-to-ship providers, authorities, take fuel samples in West Africa–those were key for our success", Marietta Harjono concludes.

In total, the research that leads up to the report took almost three years. Journalists from media outlets in Switzerland and abroad were integrated into the dissemination of the story. The impact of the campaign was impressive. As a result of the investigation, Ghana significantly lowered the limits on the sulphur content in imported fuels and other countries have put the issue on their agenda. This happened just ten months after the publication of "Dirty Diesel". Without a deep understanding of the trade business and the use of so-called data journalism techniques throughout decades, the publication of this effective report would not have been possible.

"Being able to master data will be essential for investigative work in the future", says Gian-Valentino Viredaz. Hence, there is an awareness of the potential in the multitude of datasets available and in need of scrutiny among the team members. But neither of the interviewed colleagues see themselves as frontrunners in the field of data investigations and few people have had the time or resources to dig deeper into research techniques that involve coding.

Why Are Data Analysis Techniques Not More Widely Embraced by NGOs?

The use of information technology in human rights fact-finding is not fully established for several reasons. One of them might be that NGO researchers come from different backgrounds. There is no normalised course or training to become a human rights investigator. Neither is there a common understanding of what it entails to be a human rights investigator, whereas investigative journalists more often share their skills through both national and international networks.

The organisations that conduct research in human rights violations can span from a group of indigenous people in the Brazilian rainforest to a worldwide operating network of policy analysis experts. And even within the respective associations, the differences can be enormous. As Ilia Utmelidze (2013, p. 447) from the Case Matrix Network writes: "The use of information technology in fact-finding work is often associated with empirical examples of successfully implemented database projects; but also with the disappointments, frustration and failures frequently caused by a lack of good communication between professionals of different disciplines such as law, information technology, political science, statistics, or sociology."

What Are Examples of Successfully Implemented Database Projects?

Some of the best examples of successful database projects in the NGO community are related to tracking and uncovering white collar crimes. Corruption, money laundering or fraud can have a significant influence on human rights—oftentimes these crimes are the actual source of brutal inequalities in low- and middle-income countries.

There have been many efforts to tackle the scale of white-collar crime from journalists and NGO-networks all over the world. In Eastern Europe, the Organised Crime and Corruption Reporting Project (OCCRP) employs an array of investigators, developers and designers to gain insight into public records, leaked information and source documents concerning corruption and fraud—with the goal of becoming "the people's NSA" as branded by a Freethink documentary. OCCRP has published a number of high-profile investigations since it was founded in 2006 by two investigative journalists. Among others, the initiative has revealed the largest case of tax fraud in Russian history and was involved in revealing massive money laundering schemes across the globe.

Based in the Netherlands, the organisation BankTrack (BankTrack, n.d.) publishes research reports targeting private sector commercial banks and the unjust or illegitimate activities they finance. Presenting these reports through their so-called map of "dodgy deals", makes the information available to journalists and the broad public. BankTrack was founded in 2004 as a network of twelve civil society associations. By now it has grown to include almost 40 organisations and has adapted its own status as a global, membership-based NGO in 2015. Their work is paid for by various foundations.

The Publish What You Pay (PWYP) coalition of civil society organisations (Publish What You Pay, n.d.) advocates for financial transparency in companies that are involved in mining, oil extraction and similar lines of business, commonly known as the extractive industry. PWYP has pushed extractive companies and governments to publish information on their finances and operations and uncovers the stories hidden in that data. Also, the coalition runs a "Data Extractors Programme" to train participants to uncover and make sense of extractives data. Their goal is to enable citizens all over the world to help ensure natural resources are managed for the benefit of society as a whole, according to their website.

PWYP offers training and workshops as well as online learning opportunities. The organisation also gets funded by foundations and grants.

Where It Gets Complicated

The aforementioned groups are all working with developers and open data activists to access the relevant data in order to unveil the wrongdoings of companies and politicians. There is a lot to be found in company reports and publicly available records, yet still too many databases are left undisclosed. However—in the world of business and trade, where knowledge is power, the amount of data is vast. But there is also a world beyond yearly reports and ready-made spreadsheets.

In some contexts data is scarce. In war-torn regions, in countries struggling with political turmoil or in remote areas, it is unlikely that a formal office operates to aggregate data on the atrocities and shares it through a website. This is a scenario where researchers have to go into the field and see the problems people are facing with their own eyes—and then come up with innovative technological solutions to face the scope of the human rights violations they witnessed. In the world of data journalism, it often goes unnoted, that a lot of databases have yet to be established.

In difficult and at times dangerous settings, the first challenge for human rights investigators is to find reliable long-term partners on site—or in other words, to set up the databases they want to establish on solid ground. The availability of information depends heavily on the state of the conflict and also on the international attention. In the case of the ongoing war in Eastern Ukraine, Friedhelm Weinberg, director of the Geneva-based organisation Huridocs, has observed difficulties in finding partners with a dependable financial backing. "The longer the conflict lasts, the more difficult it gets for our partners to stay involved. In the beginning, there is a lot of interest, a lot of funding, everybody wants to cooperate. But when the attention diminishes and the money stops flowing, it gets more difficult and more dangerous to obtain information."

Human rights investigations were also thriving in Georgia after the Russo-Georgian war about Abkhazia and South Ossetia in 2008. The interest in the region has since worn off dramatically, says Weinberg—Georgia is not on the human rights map of interest anymore, even though the situation of internally displaced people has not ameliorated. The attention economy of the news media affects the work of human

rights organisations which affects in turn the resources allocated and the data gathered.

Data collected in difficult contexts and with vulnerable members of society is often patchier than any data journalist would wish for. A lot of NGOs are small and thus making existence claims is often the only way they can prove a point. Such claims often do not even fill an excel table. Yet it is essential to have this kind of evidence in a human rights context. "If the ambassador of a state stands in front of the United Nations and says that there is no problem with women's rights in his country–then it is incredibly important that we can say: We have evidence on at least 50 cases, that prove that there is a problem", says Friedhelm Weinberg.

What Are the Solutions?

Huridocs (Huridocs, n.d.) was founded in the 1980s as a documentation centre for the classification of human rights violations. Soon after the launch of the organisation, a need for technological solutions emerged, aiming to make data quantifiable and searchable. Today, Huridocs offers data solutions for a variety of human rights investigation efforts. Asked to describe his work, Friedhelm Weinberg gives this example: "Let's look at a team that works on the issue of torture in Tunisia. They collect various pieces of information. They conduct interviews with witnesses, with victims, with their relatives, they collect medicinal evidence. All these pieces of information are used to push cases forward. But how do you organise this information? And how do you analyse and quantify it years later?" This is the task of his team.

Huridocs is not the only organisation that offers IT solutions for human rights investigators. Similar services are provided by Benetech (Benetech, n.d.), a non-profit based in the Silicon Valley of California and entrenched in the high tech industry that shapes the region. Another non-profit team with the same goal is The Engine Room (The Engine Room, n.d.). Their international team of policy experts and developers is fully digital—they do not run a head office, everything is organised in conference calls and via emails and other channels of digital communication. These organisations all earn their income from a range of NGO customers and also get funded by donors.

Such organisations offer information technology solutions that support human rights fact-finding every step of the way: in the actual research and data collection, data transfer, document management and

archiving, situation or fact mapping, case management, legal and statistical analysis, the provision of access to public records as well as outreach and knowledge management (Bergsmo 2013).

Examples of Successful Implementation

Precisely because finding statistical evidence in situations of conflict is so difficult, it is all the more important to pursue employing data sources. In the case of Chad, there existed numerous allegations of systematic torture and crimes against humanity concerning the rule of Hissène Habré between 1982 and 1990. When Human Rights Watch gained access to thousands of documents generated by the security forces at the time, they hired the Human Rights Data Analysis Group (Human Rights Data Analysis Group, n.d.), an organisation also based in California's bay area, to digitalise the evidence. Making the data machine-readable and thus analysable made it possible to prove that the former president was well informed of the hundreds of deaths that occurred in prisons operated by the security forces. In 2016, Habré became the first dictator convicted for human rights abuses based on universal jurisdiction (*The Economist* 2016).

Another recent example is an investigation conducted by Huridocs, together with the International Center for Advocates Against Discrimination (ICAAD). Huridocs applied automated text analysis to reveal patterns of gender discrimination in judicial sentences of sexual and gender-based violence cases in countries of the Pacific Islands. The work uncovered to what extent court decisions were influenced by gender biases and how the access to equal protection and justice was denied to women (Widmann 2017).

For some human rights organisations, the sky is not the limit when it comes to digital investigation. Human Rights Watch for example utilised satellite imagery to prove human rights abuses in an investigation of the destruction of Rohingya villages in Myanmar by comparing before and after images of affected regions. Similar work was conducted by Amnesty International for an investigation on attacks by Boko Haram militants in North Eastern Nigeria.

Other teams gather and fact-check data in ongoing conflicts in order to document human rights violations and strengthen the collective memory. Examples are the Syrian Archive (Syran Archive, n.d.), an initiative launched by a collective of Syrian human rights activists, or the Turkish

organisation Hafiza Merkezi (Hafiza Merkezi, n.d.) that does similar work in context of the current autocratic regime in Turkey.

What Is Next?

This short insight into various (digital) data sourcing techniques used by NGOs has shown that the possibilities of data-driven human rights investigations seem endless. With more people uploading information to the web, with more governments implementing open data policies, and more investigators knowledgeable in the field of data research—it becomes difficult to envision the scope of what is yet to come.

Surprisingly, for Friedhelm Weinberg, the increasing amount of data available is not only a source of joy. The larger the ocean of information becomes, the harder it will be to find the relevant pieces, he reckons. Even if it sounds contradictory: A lot of information can be hidden within transparency. The tools of human rights investigators are powerless in the face of what multinational data corporations are able to do, he claims. There is the danger, that those that already have the technology get stronger with every new piece of information, whereas those who are struggling with basic tools are further left behind.

To tackle this bias, The Engine Room, Benetech and Amnesty International have jointly published a guide on navigating digital data for human rights research. "[H]uman rights researchers appeared to be overwhelmed by the possibilities", they write in the introduction to the handbook titled DatNav (Amnesty International, Benetech, and The Engine Room 2016, p. 8). "In the face of limited resources, not knowing how to get started or whether it would be worthwhile, most people we spoke to refrained from even attempting to strengthen their work with digital data."

As mentioned above, a lot of organisations in the field have already embraced the work with digital data. And many other human rights investigators are well on their way to doing so. Something that will certainly also increase, is the cooperation between advocacy associations and journalists. As many media outlets have to cut down on their investigative budget due to fewer revenues in advertising, they increasingly have to rely on the work by human rights fact-finders more than ever before. Another encouraging development is the rise in the number of investigations conducted in larger networks, including both journalists and advocacy groups.

Although journalists and NGOs might not always have the exact same goal they can help each other to achieve more in their respective missions to uncover the truth. For journalists this means always being critical and taking the political agenda of the advocacy group into account. For NGOs it means looking for solutions to take into consideration all the possibilities at hand. In DatNav (Amnesty International, Benetech, and The Engine Room 2016, p. 9), the editors write in a personal note to the reader: "Overcome fear of digital data and technology, which is already in heavy use by your counterparts. With better tools, you know you can be more efficient."

With better tools, it will be easier to fight human rights violations across the globe.

References

Amnesty International, Benetech, and The Engine Room. (2016). *DatNav: How to Navigate Digital Data for Human Rights Research*. Retrieved from https://www.theengineroom.org/wp-content/uploads/2017/01/en-datnav-report_high-quality_web_.pdf.

BankTrack. (n.d.). [Homepage]. Retrieved from https://www.banktrack.org/.

Benetech. (n.d.). [Homepage]. Retrieved from https://benetech.org/.

Bergsmo, M. (Ed.). (2013). *Quality Control in Fact-Finding*. Retrieved from http://www.toaep.org/ps-pdf/19-bergsmo.

Hafiza Merkezi. (n.d.). [Homepage]. Retrieved from http://hafiza-merkezi.org/.

Human Rights Data Analysis Group. (n.d.). [Homepage]. Retrieved from https://hrdag.org/.

Huridocs. (n.d.). [Homepage]. Retrieved from https://www.huridocs.org/.

Public Eye. (2016). Dirty Diesel. *How Swiss Traders Flood Africa with Toxic Fuels*. Retrieved from Public Eye website: https://www.publiceye.ch/fileadmin/files/documents/Rohstoffe/DirtyDiesel/PublicEye2016_DirtyDiesel_A-Public-Eye-Investigation.pdf.

Publish What You Pay. (n.d.). [Homepage]. Retrieved from http://www.publishwhatyoupay.org/.

Robertson, G. (2010). Human Rights Fact-Finding: Some Legal and Ethical Dilemmas. *University College London Human Rights Review, 3,* 15–43. Retrieved from https://www.ucl.ac.uk/human-rights/research/ucl-hrr/docs/hrreviewissue3/robertson.

Syrian Archive. (n.d.). Syrian Archive. *Curating Visual Documentation for Justice and Human Rights*. Retrieved from https://syrianarchive.org.

The Economist. (2016, June 4). One Dictator Down. *The Economist*. Retrieved from https://www.economist.com/news/middle-east-and-africa/21699871-chads-former-president-has-been-convicted-whos-next-one-dictator-down.

The Engine Room. (n.d.). [Homepage]. Retrieved from https://www.theengineroom.org/.

Utmedlidze, I. (2013). Information Technology and Quality Control in Non-Criminal Justice Fact-Work. In M. Bergsmo (Ed.), *Quality Control in Fact-Finding*. Retrieved from http://www.toaep.org/ps-pdf/19-bergsmo.

Widmann, N. (2017, March 21). *Applying Machine Learning to Detect Judicial Bias in the Pacific Islands* [Web log post]. Retrieved from https://huridocs.org/2017/03/applying-machine-learning-to-find-judicial-bias-in-the-pacific-islands/.

Check for updates

CHAPTER 8

Following the Money Trail: Investigative Data Journalism

Miranda Patrucic

In December 2014, my friend and colleague, Khadija Ismayilova, was arrested. Khadija's work had exposed the ruling Aliyev's family control of major Azerbaijani industries including construction, mining and telecom. She also proved that President Aliyev's wife and daughter had received much of the purloined assets. Her stories earned the president's wrath. With her imprisonment, Azerbaijan was going to lose not only its best journalist, but one of the few persons investigating the first family's corruption.

Prior to her arrest, she had one request for us: continue her work. It was a powerful message to the authorities. Jailing journalists would not stop investigative stories from being published.

It was easy to assemble a team of reporters willing to tell her stories. More than two dozen reporters joined. But the actual reporting was far more difficult.

Other than through Khadija's reporting, foreign journalists knew little about the Azerbaijani ruling family. They did not speak the local language and lacked sources and contacts in the country. They teamed

M. Patrucic (✉)
Organized Crime and Corruption Reporting Project (OCCRP), Sarajevo, Bosnia-Herzegovina

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_8

up with some local reporters, but few of them had skills to investigate corruption.

Yet, by the time Khadija was released a year and half later, our team produced more than dozen top-notch investigations focusing on the first family, which uncovered hidden assets, bribery, grand corruption, and abuses of office.

For example, we uncovered how a company close to President Aliyev apparently took over, with little or no payment, the government's stake in the country's largest mobile operator, which was worth about one billion dollars. The takeover was aided by a subsidiary of TeliaSonera, the Swedish-Finnish telecom giant, which acted against its own financial interests to allow the deal. TeliaSonera facilitated and financed the takeover and agreed to dilute its own ownership stake and turn down dividends in order to placate the government. In exchange for this arrangement—essentially a massive bribe—its local partner, Azercell, would receive all the regulatory approvals and licenses it needed to operate in the country. A former financial investigator who reviewed the findings called it possibly the largest bribery case in Swedish history.

In a separate story, we also reported that members of the Aliyev family and their close advisors own significant shares in at least eight major Azerbaijan banks controlling assets worth more than $3 billion. Our reporters found that the Aliyevs own at least one-third of the five-star hotels in Baku, many built for the 2015 European Games.

In still other stories, OCCRP found a mansion and luxury apartments in London, a dacha in Moscow, and two luxury yachts that sailed the world—all owned or used by the first family. On one occasion, the Aliyevs had the 54-meter yacht sail across the Mediterranean so their 19-year-old son could have lunch in Sardinia—at a cost of more than $100,000.

Despite knowing very little about Azerbaijan when we started, years of tracking the hidden wealth of corrupt officials in other countries helped us do the same in Azerbaijan.

After all, Azerbaijan is not the first and far from the only country where officials have figured out how to embezzle the national wealth.

Former Egyptian President Hosni Mubarak, Uzbekistan's former president Islam Karimov, and Thailand's former Prime Minister Thaksin Shinawatra have similarly found ways to hide the money allegedly embezzled over their many years in office. They did so by establishing corporate structures that include companies in Cyprus, the British Virgin

Islands, Egypt, Switzerland, Panama, and the United Kingdom. They bought land and real-estate and some entered into joint businesses with the governments they themselves run.

No official statistics exist on the "proceeds of corruption"—that is, how much money is stolen by public officials from taxpayers around the world. But the World Bank estimates that developing countries lose $20–40 billion a year through corrupt acts such as bribery, embezzlement, and misappropriation of property or funds.

Secret Swiss bank accounts are out of fashion these days.

Leaders looking to abscond with profits from bribery, embezzlement, or other corruption do not want to transfer funds between bank accounts. Money transfers leave traces that investigators can track. Instead, the smartest corrupt leaders now stash their ill-gotten loot in international financial centres, where it is invested, laundered, and changed into real estate, cars, loans, boats, planes, and companies.

Unless you have a great source, bank records are close to impossible to obtain. What you can do, however, is to figure out how they spend their money. Many corrupt officials like to spend lavishly. And when their spending greatly exceeds their legitimate income, you know there is a great story there for you to find.

Here is how you do it.

How Much Do They Earn?

The first step is to find out what an official is earning.

Information about salaries paid to public service employees and government officials is a matter of public interest and can usually be obtained through a request to the institution where they work, whether that is a parliament or a particular ministry.

When filing such a request, ask for everything—not just an official's salary, but all the other benefits he or she may be receiving, including per diems, bonuses, paid transportation costs, housing costs, and so on. In some countries, you can also request information from a tax office about the earnings officials have reported, which, in addition to their salaries, should include stock dividends, fees, and all other sources of income.

Next, research the biography or curriculum vitae of the person you are backgrounding to check which boards they may be members on, or whether they are earning speaker fees.

Asset declaration forms, required of politicians in many countries, are also worth checking. They contain much more than simple earning information. Often, these declarations even include close family members.

Just do not believe everything on them. When reporters from the Center for Investigative Reporting (CIN) in Sarajevo compared earnings reported by the members of parliament with official salary records from Parliament, they found that some legislators had failed to report as much as half of their actual earnings.

But even if they are not fully accurate, asset declarations can give you leads on what the subjects of your potential story own: land, houses, cars, savings accounts. They could also tell you what they owe in the form of loans and liens, and to whom.

Tracking Assets

In many countries, land records are available online. Some can be searched only by a number under which a property is registered, while others allow name searches or let you look up personal ID numbers. If you do not have an ID number, you can ask registry staffers to do a search for you. Make sure that you at least have a date of birth, so you can distinguish among people with the same name.

Online records can also tell you if property has been mortgaged. Learning the date of a mortgage allows you to put together a timeline of when someone acquired their possessions. These records also show the amount of a mortgage, which bank gave it out and under what terms, who may have partnered to share or co-sign on a property loan, and from whom land was purchased. This mortgage information is commonly available in most countries. Make sure you also check countries your subject visits or where they take frequent vacations, as they might be making deposits and taking loans in those jurisdictions.

Do not stop there. In addition to what your subject owns, you want to find out how they got it.

Several years ago, I was part of a team that reported on the swanky apartment of Bosnia's former prime minister. The news interest in this case was that the apartment was bought with public money and then given to the official for just a fraction of its value. By following the money through property transaction documents, we were able to reconstruct what happened.

One day, the then-future prime minister checked out a beautiful apartment in a prestigious neighbourhood of Sarajevo, the capital of Bosnia, and knew he had found the place he wanted to live—a place worth far more than his salary could afford. Luckily, the public company of which he was director provided half the money he needed, and the government run by the party he belonged to gave the rest. Then, even before the government was officially registered as the owner of the apartment, it declared the apartment as not needed and allowed the prime minister to buy it out at a fraction of its value.

The story was spelled out entirely in the documents we found, including in a contract between the government and the public company, the sales contract, a request by the government to register its ownership, the government's written decision to sell it, and the Prime Minister's request to register the apartment as his.

In short, it is nearly always worth putting up with the confusion and grime of old paper storage facilities such as land registries.

Other kinds of documents you can find in land registries are loan contracts, payment orders, and receipts that show whether payments were made in cash or by wire transfer, and who made them. In these documents, you can also see a date of purchase and the identity of a property's previous owners, as well as whether it has been used as collateral for some other purchase.

Other valuable sources of information on real estate include municipal government or other authority offices that issue construction permits, plus tax offices and real-estate agents.

Other assets that can be tracked include cars, yachts, and planes. In many countries, information on who owns these is available online. A registry of mortgages can be useful. If a person has bought a plane or a car with a loan, those assets will be mortgaged and record of that is kept. You can ask for original documentation as well, including sales contracts to check the value of assets.

Tracking Companies

Most countries now allow access to company records online. But in many cases, these online databases cannot be searched by the name of a company's owner, only by the name or ID of the company. If you do not know the name of a company a person of interest owns, you can try requesting it from the official registry. Online databases are now

available that allow for searches across multiple databases, the best being LexisNexis, Dialog, and Dun & Bradstreet. These databases offer endless possibilities when it comes to researching people and companies, including doing keyword-based searches across a number of databases that can reveal a network of companies. Reporters abroad can often get free access to these resources at diplomatic libraries, such as at a US Embassy. They can also usually be found at public libraries or university libraries.

Websites like OCCRP's Investigative Dashboard (ID) and ICIJ's OffshoreLeaks offer search tools that give reporters access to large amounts of data brought together from online databases and leaks. On ID Search, journalists can already search information from over 200 sources, including many leaks and government databases published in recent years.

When business databases are not searchable online, you can try searching official gazettes. In many countries, companies are obliged to publish notice of incorporation, of change of directors, and shareholders' addresses in these publications. Try searching through all editions of gazettes for the people and companies you are interested in.

When looking up a person's name, do not give up if you do not find a company at first glance. Always try variations, like searching only a surname or different spellings. To avoid identification, people often purposely misspell their name, use a middle name or maiden name, or just make up a name. In some cases, business registers look the other way. For example, Khaled Salam, former director general of the Palestine Investment Fund and economic adviser to former Palestinian President Yasser Arafat— who was recently sentenced in absentia to 15 years in prison and a $15 million fine for embezzlement and money laundering— used the name Mohammad Rashid when he set up a company to buy coastal property in Montenegro.

You should also try searching by both personal and company addresses. In some cases, you might find multiple companies registered at the same place. A number of companies connected with the family of Azerbaijan President Ilham Aliyev were found registered all at the same place: the family's home address.

To hide their identity, people often use a proxy—a person who knowingly or unknowingly acts on behalf of another. This may be a close relative, spouse, or child but could also be a business partner, friend, trusted lawyer, or employee. When looking up companies, always check additional names, and also run a search by the names of their registration

agents. People in the same circle often use the same agent to establish companies.

If someone wants to keep his ownership truly hidden, they will establish multiple layers of secrecy. They can incorporate a company in the British Virgin Islands, which in turn will set up a company in Belize, and from Belize a company in Sweden. It is important to follow all these leads and try to get names and addresses of the proxies for all these companies.

What online records cannot tell, documents in a business registry might. When an offshore company incorporates a local company, it needs to authorise someone to do it. That gives you a name that may even be a member of the family or, rarely, the person you are looking for himself. Miroslav Miskovic, a Serbian tycoon who controls 6 percent of that country's GDP, used his employees to set up offshore companies and buy land in Serbia for him.

Make sure you go back in time, too. In some databases, such as on LexisNexis, only the current information is available. But if you look at old records, you might find the connection you are looking for. Archival data is often available. Get the company's whole history.

Obtaining hard copies of company records is also important—they can hide treasures. You can find all sorts of documents—from loans, to partnership agreements, to court verdicts. These, looked at by too few reporters who are intimidated by their complexity, can yield copies of IDs, passports, or registration records of their owners. Always get everything you can. You never know when a piece of information out of a document will prove valuable.

Finally, get financial reports and audits. In most countries, these are public documents. They will contain information about connected companies, transactions with related parties, including owners, and if you know how to read them or can find someone to show you, they will tell you all you need to know about business activities.

Above all, do not give up and keep digging. A great story is waiting around the corner.

Tracking Offshores

Offshore destinations are a favourite choice of those seeking to hide assets, launder money, or evade taxes. In addition to confidentiality and flexible incorporation procedures, they make only minimal demands for the disclosure of corporate information.

In the summer of 2011, I first learned about Gunvor Group. After months of intensively investigating one of the most successful oil giants and its operations, a colleague from Uppdrag granskning, a Swedish television program, asked me this: "Is Gunvor transparent? Yes or no?"

His investigation was prompted by the Wikileaks release of US State Department cables alleging that Russian President Vladimir Putin had a stake in the company. The US Ambassador to Russia said that Gunvor had close connections to the Russian government and that its secretive ownership was rumoured to include Putin.

In an interview for Uppdrag granskning, Torbjörn Törnqvist, a Swedish oil billionaire, refuted the allegation and claimed that he and Gennady Timchenko, a Russian tycoon with Finnish citizenship, each own about 45% of the company, while 50 employees own the rest.

The official records told a different story: The Gunvor Group was owned by a series of nested structures made up of trusts and offshore companies that were specifically designed to prevent anyone from knowing who actually owned anything. Sven Olsson, a lawyer from Helsingborg who for many years has been a central figure in Timchenko Swedish and international companies, was the only person who could be identified.

Structures like these are used for three major reasons: to avoid paying taxes, to launder money, or to keep the authorities from finding out the real owners. That Törnqvist and Timchenko actively hid who else was involved in the other parts of the company meant that Putin could very well have a stake.

A few months later, in Montenegro, as I was showing a group of reporters and anti-corruption activists how to access Cyprus business records, one asked, "Can we check who owns Numerico Trading Limited?" This company had signed a deal with local authorities to build the first aqua park in the Balkans and got them to guarantee a €1.6 million loan. The park was never built, and the Montenegrin community of Budva was left to foot the bill. For years, the only thing known about the company was what local authorities had claimed at the beginning—that it was a reputable Spanish company.

I searched through the Cyprus business registry and cheered as I opened a list of shareholders. The owner of the company turned to be Milos Marovic, son of the former president and one of the country's most influential politicians.

Tracking offshores is not a "Mission Impossible". You might not get the full results you want, but then again, you just might. You do not know unless you systematically track through available databases. It is not easy and it takes patience and determination, but sometimes you get lucky.

Until a few years ago, getting records from offshore zones was difficult and expensive. Lately, some of these jurisdictions, including Panama, Cyprus, and the Isle of Man, have opened up websites where records are obtainable for free or cheap. Others—zones like the British Virgin Islands, Belize and the Seychelles—require written requests to the registry. While they are more open, actually getting records require some tricks we are slowly learning and sharing.

The Future

This technology is poised to take a giant leap, where search tools will start more active roles as "assistant researchers." New knowledge management databases can help to figure out all the places where a particular politician or company has been mentioned across many different databases, automatically generating maps of the asset ownership of ruling families and other elites. Machine learning can help you separate common names and help give you more accurate results. Artificial intelligence can use past patterns of criminal structures to find yet undiscovered criminal activity. Already today we can upload a list of members of a powerful family and link their profiles to relevant information from Wikipedia for information about their family structure, to the Panama Papers for offshore assets, land registries, company registries or leaked email inboxes from a neighbouring country's politician. Soon the systems will tell us where to look first for criminality. Journalism assistance bots will tell you what is happening in the world and what you might be interested in.

Reuters' Head of Data and Innovation, Reg Chua, has described this as "structured journalism": Research findings from human researchers will increasingly be recorded as machine-readable data that can be augmented and cross-referenced by algorithms that encode notions of journalistic relevance. The outcome will also facilitate sharing amongst journalists, hinting to journalists when they need to cooperate to crack a case across borders.

The documents leaked by NSA whistleblower Edward Snowden have demonstrated the extent to which modern government intelligence work is based on bulk data analysis. Investigative reporters must make sure they leverage digital technology to hold the powerful to account, as well.

PART II

Immersive Storytelling and Visual Investigation

Over the past years, experimental and progressive approaches of journalistic investigation have emerged. Continuous digitalisation has created a gap between information formats and journalistic tools or skills that are required for processing such information. To bridge this gap, journalism needs to develop approaches that encompass other expertise aside from traditional reporting. Within technology-driven projects, we can observe highly differentiated teams that include, amongst others, technologists, computer scientists, architects, designers, and visual journalists. As such journalism platforms assemble teams of diverse practitioners, also the profile of a journalist is undergoing changes. Journalists need to enrich their professional repertoire to meet the specific requirements of digital investigation.

Particularly digital audiovisual content poses new challenges for journalism. The anonymous publication of videos via social media is ubiquitous, their viral and user-based dissemination is rapid. As a result, identifying originators of such content is demanding. It can be only achieved with the help of special techniques such as crowdsourcing and visual analytics, for instance.

These new techniques and investigative approaches push the boundaries of journalism, ultimately advancing it and creating new genre-spanning concepts. However, traditional journalism remains the pillar of news production. But due to its organisational structures and economic obligations, traditional journalism can only—if at all—slowly adapt to changes and implement innovative and oftentimes risky concepts.

Experimental and unorthodox methods are either applied only on a small scale or on highly specialised niche platforms. In the following chapters, contemporary approaches of storytelling and contextualised investigation as well as accompanying practices of data verification and authentication are portrayed by contributors who represent a constantly changing journalism.

Check for updates

CHAPTER 9

Drone Journalism: Storytelling from a New Perspective

Ben Kreimer

Introduction

Drone technology has democratised access to previously inaccessible aerial vantage points, spaces, and perspectives, while opening up new data gathering, sensory, and spatial documentation possibilities for journalists. Functioning as a much safer, more cost-effective, accessible, and easier to operate alternative to manned and analog remote-control aircraft, aerial drones create new storytelling possibilities from the air.

This democratisation also provides, for the first time, a way for journalists to gain uncensored access to aerial perspectives in countries where, in the past, government aircraft have been the only way to get a reporter's camera in the sky, a point made by Dickens Olewe, the founder of African skyCAM: Africa's First Drone Journalism Team. Olewe started African skyCAM when he won an African News Innovation Challenge grant from Code for Africa in 2012 while working as a journalist at The Star, a daily newspaper in Nairobi. Olewe has said that disaster reporting is one area where drones can help journalists,

B. Kreimer (✉)
Lincoln, NE, USA

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_9

especially in Global South countries like Kenya, where reporters may resort to riding in government aircraft to capture visuals of the damage and recovery efforts. Although journalists are allowed onboard the government aircraft, the pilots, and their overseers, have the final say on where they will and will not fly, putting journalists at the mercy of governments that may be at odds with journalism.

Aside from replacing manned aircraft, drones can also operate in ways that airplanes and helicopters cannot. In the United States, manned aircraft are normally required to stay above 500 feet, except when landing. Drones, on the other hand, can safely and legally operate from ground level up to 400 feet. This dynamic range of operation at varying altitudes, made possible by their small size, makes drones a storytelling tool that can gather data, images, and videos that enlighten stories by taking audiences into inaccessible spaces, capturing surprising visuals, and presenting the world from a new and different perspective.

When to Use a Drone

Drones are essentially flying robots given their onboard sensors and flight controllers. As the term drone suggests, they are capable of varying levels of autonomy and come in different shapes and sizes. Although this chapter focuses on flying drones, there are also land-based and underwater drones that function similarly to their aerial counterparts, leveraging sensors and onboard computational power for varying degrees of human control and autonomy during operations in their respective environments. Given their range of capabilities, journalists should carefully consider how a drone can enlighten a story and captivate audiences. Simply because a drone is accessible does not mean it is the best tool for the job, especially when considering how the novelty of technology can be a distraction from the focus of the story. Yet, that novelty factor can also enlighten the story, hence the need to always ask: "Why use a drone for this story?"

Drones are a great tool to use when a story has important spatial and scale components. The images and video captured by a drone can add critical context to a story in ways ground-based visuals fall short. As with any other image or video, the storyteller should consider the effect of the drone visuals on the audience, and how they connect to the other content used to tell the story. While drones can add to a story by revealing a radically different perspective than what's observed on the ground, they

can also distort reality for the audience, sometimes in unexpected ways. This can present both opportunities and challenges for journalists striving to report on events and conditions as accurately as possible.

For example, while working with African skyCAM on a video documentary about Nairobi's Dandora Municipal Dump Site, a controversial overflowing and toxic landfill known as a dumping ground for biological and pharmaceutical waste, Dickens Olewe commented on how, ironically, the landfill looked beautiful from the drone captured images and videos I had taken from high above, which are viewable on my website (Kreimer 2014).

Given that the underlying reporting and ground-based visuals highlighted how toxic the landfill is to people who work there and live nearby, the drone visuals alone fail to convey the human story, although they do provide an important perspective, showing the scale of the landfill, which complemented the story we were capturing on the ground about the people who survive off of the materials they find in the landfill and then sell. If the only visuals we had were captured by drone, then I would have reconsidered using them in the story, given the lack of ground context, which the aerial perspective should complement, and vice versa.

Because the drone functions as an airborne extension of the journalist, it is important to remain open to surprises. Humans live on land, and our understanding of the world around us has been largely shaped by experiences on the ground. So drone journalists should prepare themselves for visual and perspective surprises when flying a drone. It is wise to do research before flying, not only to check for legal and spatial information (airspace, nearby airports, et cetera) and ground hazards but also to get a sense of the story you are going to tell with the drone. Look at satellite imagery, photos, 3D buildings on Google Earth, and other information regarding the site. Still, when you're flying the drone, expect visual surprises and opportunities, and be prepared to take advantage of them.

Speaking of surprises, given the aerial perspective and unique view that camera-equipped drones can provide of the ground, the technology is also capable of revealing stories. In December 2011, an amateur drone operator in Texas spotted a blood red creek outside Dallas. As reported by the Columbia Journalism Review (Hepworth 2016), the operator had gone to test a drone, coincidentally flying near the red creek that ran from behind a factory, and into the Trinity River. Upon making this discovery at home while reviewing recorded aerial video, the unnamed

operator called the National Response Center, which then communicated the find to the Texas Environmental Crimes Task Force. The factory was a slaughterhouse and meatpacking plant that had been dumping pig blood into the creek. Next came a search warrant, a court case, and 18 indictments against the plant and two of its vice presidents.

Two years later, the Texas Legislature created the 2013 Texas Privacy Act, which clamped down on drone use. A drone operator may now be fined $5000 per image for capturing, without permission, people or private property. Had the Texas Privacy Act been in place in 2011, the drone operator who photographed the blood red creek would have been subject to legal action by the meatpacking plant because the video clips were captured without permission. This exemplifies the intersection of drones, their ability to expose stories, and how rules and regulations come into play.

Rules and Regulations

Drone rules and regulations range from no regulations at all to outright bans that prohibit civilian drone use. Many countries now have nuanced rules and regulations that spell out operating conditions that must be adhered to during drone operations, particularly for commercial use, which in the United States, UK, and other countries includes journalism. For countries with nuanced drone regulations, certification requirements for commercial operators are common. Still, the information in this chapter is merely a stepping stone for finding the information you would need to seek out prior to conducting legal drone operations.

Drone rules and regulations vary widely and can come from communities on up to national governments. Rules and regulations are also evolving over time and will continue to do so, following the arc of drone technological innovation, the reliability and safety of the hardware and software, combined with public sentiment and government comprehension and support of the technology. As drones evolve and become increasingly autonomous, more reliable, and safer to use, rules and regulations will likely also evolve, albeit more slowly than the pace of technological innovation.

Rules and regulations are expansive and dictate what you can and cannot do with a drone. In many places, drones are not allowed to fly over 400 feet above ground. Flying over or near people is also regulated, to varying degrees from country to country. There are also regulations that

state how close you can fly to wildlife, factory farms, buildings, airports, and critical infrastructure. As is the case in Texas, there may also be personal and property privacy elements to be aware of. Rules also exist regarding nighttime operations. You may legally be required to equip your drone with special lighting if you are flying before sunrise or after sunset. This by no means exhausts the range of rules and regulations for operating drones, which is why operators must do their research before flying if they desire to operate legally.

Prior to flying, a drone journalist should find answers to these questions:

- Am I permitted to fly a drone in this community, state, and country?
- Will the drone operation occur before sunrise or after sunset?
- Will I be flying around people, and how close will they be to the drone?
- Who owns the property I am flying over, or near to?
- Can I fly on that property, or do I need to get permission to do so?
- What is the airspace, and how close am I to an airport?

These questions are an analysis of your operation and the airspace you plan to operate in. Check the current rules and regulations and airspace restrictions for your operation location. Some drone rules and regulations, if followed literally, may go against the principles of journalism, especially when it comes to where you can and cannot fly. For this reason, and because drone technology and regulations are new and continue to evolve, there are legal grey areas. Seeking legal counsel is advised if you are unsure about the legal status of your operation.

Ethics

Drones are an airborne extension of the storyteller, providing access to airspace locations that may otherwise be out of reach. When looking for an ethical framework for drone journalism, look online for the Society of Professional Journalists (SPJ) Code of Ethics and the National Press Photographers Association (NPPA) ethical guidelines, which are particularly useful considering the widespread use of drones for visual journalism.

In short, do not do something in the air with a drone that you would not do the equivalent of when reporting on the ground. For example, if you would not climb a fence to take photographs over the top then do not use a drone to do the same. The NPPA's (2016) guideline of "treating all subjects with respect and dignity" is also noteworthy given most drones are not tethered to their operator. While there may be temporary flight restrictions placed around disasters, fires, and other such events, the operator and their own ethical standards may determine whether or not a drone flight happens, or how it is performed. Drones allow journalists to engage with an event from a distance. In some cases, such as a war zone, this could keep the reporter safe. In other situations, such as if a family's home is on fire, a journalist could launch the drone from half a block away and capture photos of the family as they stand on their lawn, watching their house burn to the ground. Not only could this be considered an intrusion on the family, but such a drone operation could also "alter events," in this case obstructing the work of the firefighters fighting the blaze.

The drone and privacy debate is contentious and subject to debate in various circles, from all levels of government to online message boards. Drones can fly well beyond the operator's line of sight, making it easy to feel detached and less responsible for the drone's actions, while simultaneously making it a capable tool for recording visuals from vantage points where the drone operator goes unnoticed. It is common for operators to fly over sensitive areas and escape prosecution, as Al Jazeera (2015) reported in a story about unidentified drones flying over Paris landmarks and nuclear power plants. In other cases, blatant misuse of drones has resulted in government action and the outright ban of drones and led to operators needing government permission before flying in the country.

Before the April 2015 earthquake that devastated Kathmandu and surrounding areas, Nepal had relaxed drone regulations. That changed when humanitarian organisations, media companies, and other foreign drone operators showed up to document the devastation from above. While the initial reaction was positive, the drone operations began to cause problems. These "cowboy" operators, as Andrew Schroeder of the humanitarian robotics organisation WeRobotics, and Faine Greenwood of the SIGNAL Program at Harvard, call them, came in with good intentions, but ended up causing problems (Riley 2016). The cowboys flew over sensitive structures and religious and heritage sites that cannot

be overflown by drones without permission from the government. They also did not secure or effectively share their data effectively with humanitarian response teams, rendering it useless for recovery efforts. The cowboys also invaded people's privacy, which led Nepali citizens to complain to the Civil Aviation Authority of Nepal, who responded by requiring all civilian drone operations to obtain permission before flying anywhere in Nepal (sUAS News 2015). Had the journalism and humanitarian drone operators respected pre-existing permissions regarding the drone flights over religious and historic sites and not disturbed Nepalis, the country likely would not have come down so hard with new rules and regulations.

Multi-rotor and Fixed-Wing Drones

Drones come in different shapes and sizes but generally, fall under the categories of multi-rotor or fixed-wing. Multi-rotors come with three or more motors and propellers and are designed for vertical take-off and hovering, making them highly manoeuvrable and capable of smooth, subtle movements. The downside of the multi-rotor drone is poor efficiency due to lift coming only from the motors and propellers. On the other hand, fixed-wing drones fly like an airplane, generating lift using their wings, and can fly with a single motor and propeller, enabling them to stay aloft significantly longer than multi-rotor drones. There are also vertical take-off and landing drones, or VTOLs, that can take-off and hover like a multi-rotor drone, but also fly long distances like a fixed-wing drone.

Despite their poor energy efficiency, multi-rotors are, at the time of writing, the most common civilian drone given their manoeuvrability, ease of operation, small size, and ability to carry gimbal and camera payloads for capturing stabilised videos. Fixed-wing platforms, which have the aerodynamics of an airplane, have a minimum airspeed they must sustain to remain aloft, limiting their ability to capture the dynamic range of moving visuals that are possible with a multi-rotor platform. Fixed-wing drones need significantly more open airspace to operate, limiting their use to wide open areas, or to flying over structures and objects which multi-rotors could easily navigate through at low altitude. For fixed-wing drones, cameras are often positioned in the fuselage with the lens facing downward, making these drones ideal for mapping large areas because when operating flight time and flight distance are a priority, which is

often the case when mapping landscapes, they far outperform multi-rotors, which in most cases have about 50–75% less flight time.

Regardless of your drone, it is important to adopt a flight operations manual, including pre-flight and post-flight checklists, and maintenance logs, to ensure safe operations. I helped develop the Drone Journalism Lab's (2016) open source operations manual and maintenance logs, which are free to use and accessible online. There are also mobile device apps that assist with these processes.

Safety is important when it comes to operating drones. Responsible and safe drone use will likely lead to more friendly rules and regulations and a bright future for drone journalism. And it is also important for the drone operator and the company or organisation they work for, or were hired to work for, to maintain high safety standards and procedures. If an operator is unsafe, and their equipment is not well maintained, they will likely experience technical problems, which could lead to accidents and an unfavourable reputation, not to mention damaging the reputation of fellow drone journalists and the wider civilian drone industry.

A Flying Computer

Drones are flying computers. The value of the technology comes from the task it performs. Drones, like personal computers, are capable of complex assignments. An operator using a drone to capture an aerial photo is like an individual using a laptop for word processing. Drones and personal computers are indispensable when it comes to executing their respective rudimentary tasks, yet they are also capable of more complicated work.

Drones carry electronics that make them far easier to fly than analog remote control airplanes and helicopters. The drone's onboard flight controller is coupled to positioning sensors, which can include a gyroscope, accelerometer, magnetometer, Global Positioning Systems (GPS, and Global Navigation Satellite System (GLONASS)) satellite navigation modules. These sensors enable drones to fly themselves. When flying outdoors and with an unobstructed sky, a GPS enabled drone with adequate GPS connectivity can fly itself. This means the operator can release the control sticks and the drone will hover in place. It also allows the operator to use computer software or a mobile device app to program autonomous flight instructions onto the drone's flight controller.

This autonomy creates many possibilities for storytellers. A common function is to program a drone to fly a particular flight path free of inputs from the operator, although the operator should always stand ready to override the autonomous flight instructions in the event of a problem. This makes it simple to survey areas, capturing images or videos at preset times, or intervals, such as when the drone captures a photo every second, or at specific locations. Such drone use can provide orthorectified aerial maps and three-dimensional (3D) reconstructions when the images are captured using the photogrammetry process, which involves capturing images with significant overlap.

Autonomous flight plans also enable reproducible operations, giving the storyteller the ability to gather data over an area from approximately the same perspective and aerial location each time, opening up possibilities for time-lapse image and video sequences, and monitoring structural or landscape change over time. A flight plan could range in complexity from a single waypoint in the air to hundreds of waypoints scattered over miles. Similar techniques are used for creating aerial panoramas, whereby the operator can decide the parameters, such as degrees of coverage and the number of photos to capture, after which the drone will autonomously execute the panorama imaging sequence.

Ultimately, the full capabilities of a drone are only realised when it is paired with computer software or mobile device apps used by the operator to control or program the drone. Many companies that offer apps used to operate drones also have a software or browser and cloud computing service for data processing and analysis. These interwoven hardware and software systems are used across industries, a fact that underscores the data gathering and storytelling capabilities of drone technology, and the network of software, app, and processing tools that drone operators can leverage.

Drones for Immersive 360° Video and Virtual Reality Storytelling

With their ability to nimbly navigate airspace and carry a payload, drones are well suited for producing immersive 360° video and 3D virtual reality (VR) and augmented (AR) content. Drones can carry 360° video cameras and capture images that, through the photogrammetry process, result in 3D models and environments for virtual reality goggles, like the Oculus Rift and

HTC Vive, which are capable of room-scale VR, providing virtual experiences that enable the wearer to physically explore a space. Examples of this content can be found on my website (Kreimer, n.d.).

As BuzzFeed's Beta Fellow in the Open Lab for Journalism, Technology, and the Arts, I explored the potential of customising low cost multi-rotor drones, specifically the DJI Phantom 2, as platforms for 360° video cameras. To demonstrate this merging of emerging technologies, I co-produced a BuzzFeed News story about the aftermath of California's Valley Fire, a 76,000 acres forest fire that burned north of the Bay Area in September 2015. Given the large spatial nature of the story, the drone and 360° video camera enabled us to take the viewer on an immersive experience of the fire's aftermath, while giving us access to burned out structures that were unsafe for a reporter to traverse, demonstrating how drones can effectively navigate spaces too dangerous for people, while also capturing unique aerial perspectives. Released on Facebook in November 2015, "California's Valley Fire" was BuzzFeed's (2015) first 360° video and has, as of this writing, received over 6.3 million views. As drones and 360° video cameras have improved, capturing such content has become easier through modifying newly available drones. Yet at the time of this writing, there is not a ready-to-fly 360° video drone on the market, a reminder that hacking and modifying readily available hardware is one way of staying on the forefront of storytelling innovation. For this reason, I have made my 360° drone modifications available on my website.

Drones are also a useful tool for 3D VR and AR journalism. The use of a drone, when combined with aerial photos taken using the photogrammetry process, allows for large-scale 3D reconstructions, making it possible to bring physical spaces and objects from the real world into VR and AR experiences. The process of capturing overlapping images and processing them in photogrammetry software that understands the overlap and calculates distances between common points from image to image results in 3D reconstruction models and aerial maps that look like satellite imagery, but at higher resolutions and without interfering atmospheric clouds. The photogrammetric 3D models produced from images captured by a drone can include objects, structures, and landscape features, all of which can go into 3D VR and AR experiences.

Finding Inspiration

As drone technology evolves, its capabilities and use cases continue to grow. This is evidenced by the myriad industries and fields that are leveraging drones. From agriculture to archaeology, wildlife conservation to volcanology, to package delivery and pickup, search and rescue, and infrastructure inspection, drones are democratising access to the air. Yet, for journalists, drone hardware and software developments, and how to leverage them in the context of reporting, can remain elusive without examples that illustrate the tools at work.

One way for journalists to find innovative ways to tell stories with drones is to observe how other industries and professions are applying the technology. This kind of interdisciplinary observation and learning is crucial for understanding the capabilities of the technology. Journalists, farmers, and archaeologists are using the same drones and sensors, yet the way each industry leverages the technology differs. Observing these differences will lead to interdisciplinary learnings and innovation in drone journalism.

In 2014, I began working with a team of archaeologists in southern Turkey. Before joining the project, I learned surveying techniques and how archaeologists had been using balloons, kites, and drones, to do photogrammetry for aerial mapping and creating 3D reconstructions of landscapes and excavated structures. At the time, photogrammetry was not in use by journalists. Spotting an opportunity, after learning the process and using it in Turkey at the archaeological site, I then went to Kenya, where I used the same process to create a detailed map and 3D environment of Nairobi's Dandora landfill for browser interaction. That project, which can be seen on my website, demonstrates how journalists can do innovative storytelling by leveraging the tools and workflows of other professions and industries.

Conclusion

Drones are an important tool in the modern and future journalist and storyteller's toolkit. They make it easier than ever before to gain access to aerial vantage points that had been previously out of reach. As the technology continues to develop, the storytelling applications will continue to evolve, giving journalists even more ways of producing novel content that enlightens stories. Whether it's providing a new sense of

scale, perspective, access, or simply keeping reporters out of harm's way, drones are changing the way stories are told.

References

Al Jazeera. (2015, February 24). Unidentified Drones Seen Flying Over Paris Landmarks. *AlJazeera*. http://www.aljazeera.com/news/2015/02/unidentified-drones-flying-paris-landmarks-150224111511873.html. Accessed 20 July 2017.

BuzzFeed. (2015, November 12). *California's Valley Fire* [Video file]. https://www.facebook.com/BuzzFeedVideo/videos/1869492456524961/. Accessed 26 July 2017.

Drone Journalism Lab. (2016, August). *The Drone Journalism Lab Operations Manual* [Blog post]. http://www.dronejournalismlab.org/manual. Accessed 26 July 2017.

Hepworth, S. (2016, Fall/Winter). The Drone Files. *Columbia Journalism Review*. https://www.cjr.org/special_report/drones_photos_essay_faa_laws.php. Accessed 26 July 2017.

Kreimer, B. (n.d.). [Homepage]. http://benkreimer.com/.

Kreimer, B. (2014, October). *Nairobi's Dandora Dumpsite* [Blog post]. http://benkreimer.com/aerial-video-photography/nairobi-dandora-dumpsite-video-3d-model/. Accessed 22 July 2017.

National Press Photographer's Association. (2016). *Code of Ethics*. https://nppa.org/code-of-ethics. Accessed 28 July 2017.

Riley, T. (2016, August 25). Drones Get a Bad Rap. *Inverse Innovation*. https://www.inverse.com/article/20210-humanitarian-aid-drones-italy-earthquake. Accessed 26 July 2017.

sUAS News. (2015, May 7). Nepal Moves to Limit Drone Flights Following Earthquake. *sUAS News*. https://www.suasnews.com/2015/05/nepal-moves-to-limit-drone-flights-following-earthquake/. Accessed 30 July 2017.

Check for updates

CHAPTER 10

Immersive Journalism: How Virtual Reality Impacts Investigative Storytelling

Marcus Bösch, Stephan Gensch and Linda Rath-Wiggins

New Medium, New Stories

In 2007, American journalist and documentary filmmaker Nonny de la Peña teamed up with artist Peggy Weil to develop an immersive docu-game called "Gone Gitmo" (GONE GITMO 2007): a virtual but accessible version of the Guantanamo Bay prison camp established in 2002. Since journalists were not allowed inside the camp, the reconstruction was based on interviews with former prisoners, videos and photos from a variety of official sources and legal material. Originally designed for Second Life, the experience was later made available as a stand-alone program using the game engine Unity. Nonny de la Peña describes her work as immersive journalism: "… the use of virtual reality and 3-D environments, built in a gaming platform, to convey the sights, sounds and feelings of news" (McCartney 2011).

Ten years later, this work process still seems rather unusual for most journalists. However, the advantages are quite obvious to de la Peña and other VR journalists: Immersive technologies offer users a sense of "being there" (Maschio 2017), they allow for a very personal interactive

M. Bösch (✉) · S. Gensch · L. Rath-Wiggins
Vragments GmbH, Berlin, Germany

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_10

multimedia-experience which at the moment proves to be impressive to most viewers (Maschio 2017).

Reconstructing a scene in Virtual Reality (VR) can either be the result of an investigation (as in de la Peña's case) or the beginning, may it be for forensic reasons. Scientists from the Institute of Forensic Medicine in Zurich have been developing virtual reality technologies to recreate crime scenes (Beckett 2015). This so-called Forensic Holodeck combines information such as 3D room scans, CCTV footage and autopsy data to produce accurate 3D reconstructions of crimes. A similar approach in the context of journalism lets the user walk around the Ferguson shooting scene, where Michael Brown was shot dead by a police officer on 9 August 2014, in Ferguson, Missouri (Fusion 2015).

The danger is that the new medium focuses on characteristics already established in existing media instead of building on what is truly new, argues Helen Papagiannis in her current book "Augmented Human" (Papagiannis 2017). So, what is there to do in order to drive the medium forward without replicating what has come before? Rebecca Rouse suggests to get inspired by the past. Together with Evan Barba, she has recently published a paper on "Design for Emerging Media". They state that "[the] possibilities are still fairly open, as we continue to navigate the frontier" (Rouse and Barba 2017).

Production and Storytelling Challenges

For now, these virtual reality experiences prove to be way costlier than a written report. In addition to that, market penetration of VR devices for end users is still comparably low (Bogle 2017). But besides building complex computer-generated 3D-scenes, the advent of cheap consumer 360° cameras like the Insta Nano (Insta360, n.d.) or the Ricoh Theta allows media companies to try-out 360° experiences cost-effectively. For example, the New York Times has experimented with that kind of 360° video storytelling on a daily basis for over a year now (*New York Times* 2016). Whether journalists decide to produce 360° videos or they go for a full-blown computer-generated-based VR story, both versions are immersive for users and there is still much to learn with regard to storytelling techniques.

By now, the term storyliving (Maschio 2017) seems to be widely used in direct contrast to the traditional concept of storytelling. The reason for that lies in the immersive character of VR.

Core innovation in VR storytelling is the fact that users are at the centre of the narrative, quite literally speaking. In a 360° video production, users will take over the camera's positions. New kinds of questions have to be answered by journalists: Who is the user within that story? How do you design the presence of the viewer? To what extent does that have an impact on the story? These questions prove that it is a whole new craft that still requires experimentation and large-scale analysis of user behaviour.

According to Dolan and Parets (2016), producers have to think in two dimensions when crafting the character of the user. The first dimension constitutes the role of the users. They can be part of the narrative ("Participant"). In these cases, protagonists can talk directly to users as in the case of a 360° story we will discuss later, called "Disfellowshipped" (Bundy 2016). Contrary to that, users in a passive mode ("Observant") can observe the spherical scenes without taking over a character within the story. Both versions come with advantages and shortcomings. The second dimension constitutes the question of influence within the narration. Two modes are possible. The first one encompasses users having an impact on either the outcome of the story or the pace of the story or the sequence of the story (one example is a VR project we produced about Stasi interrogation methods (Deutschland Radio Kultur 2017)—the user has an impact on the pace of the story by having to gaze at a phone and trigger the continuation of the story). The second option is having no impact on the story whatsoever—these are examples of linear immersive VR stories (e.g. 360° videos on YouTube).

Another challenge of VR storytelling is the lack of frames. In traditional visual media, journalists had to think of constraints in terms of a 2D TV screen, a laptop or a mobile phone. In VR, these frames no longer exist as the user can turn around and look at wherever he or she wants. That raises the question of how to direct the attention to things and/or places within a sphere. How can a journalist make sure that users will look at exactly the direction that they intended to (spoiler alert: They cannot)? It also means the traditional understanding of the composition of a scene needs to be reconsidered. Zooming-in, zooming-out of objects or the sequencing and edits—many of the traditional cinematographic techniques have to be re-evaluated. Even in production mode, new challenges emerge. For example, while shooting, the production team needs to make sure that all the equipment (e.g., lighting)

is out of the picture. Also, the production team needs to get out of the way (unless their being in the picture is part of the story).

Looking at these storytelling challenges, combined with a highly dynamic VR market, journalists have to tackle a lot of different factors if they want to produce a virtual reality story (either in 360° or computer-generated).

Use Case: Disfellowshipped by Reveal

This is where our VR platform Fader comes in. We wanted to build a product for journalists to create and distribute immersive stories easily and fast without having to think about too many technical hurdles. The development of Fader was further advanced due to a cooperation with the US-based Center for Investigative Reporting (CIR) (*Reveal News* 2017). The reporter Trey Bundy was investigating child sex abuse cases within the Jehovah's Witnesses community (Bundy 2015) and he wanted to offer a more intimate understanding of a protagonist's character and her individual experiences. Bundy wanted to know whether through immersive storytelling and the new opportunities with virtual reality technologies, he can bring another layer to the story.

He approached us in 2015 and we decided to take this case for learning more about his journalistic needs while producing a 360° experience based on his investigations. This approach turned out to be highly fruitful as we were able to contribute with technical know-how and he was able to focus on the editorial part of producing a 360° story for the first time. In one of our meeting, Bundy said: "I feel like I'm writing a script for a movie without ever having been to the movie theater" (Rath-Wiggins 2016).

This shows how VR is not just a new technology but actually constitutes a completely new medium. In Bundy's 360° story, he wanted to portray the protagonist, Debbie, and her story. But how does he turn her experience into an immersive narrative? What rooms and environments does he highlight? What character will the VR users have? Will they be active? Or passive? Will they have an influence on the story? Or will they have no impact? And, also not trivial, how do we manage this project with two teams separated by an ocean and many time zones?

Our project management was comparable to many international projects: cloud-based services, Skype calls at non-conventional times (super late for the Berliners and super early for the Californians) and a

lot of emailing helped us get through the project. More importantly, Trey Bundy visited Berlin multiple times. During these sprint weeks in Germany, we were all able to initiate the concept of the immersive narrative.

Trey Bundy writes: "Post-production was a months-long trans-Atlantic endeavour. I worked with Reveal senior supervising editor David Ritsher to refine and edit the segments. Then we sent our work to Berlin, where the Vragments team integrated the elements – video, photos, animation, dialogue, ambient sound and music – and turned them into VR. The result is an experience we hope helps open up the possibilities of VR storytelling and gives our audience a more powerful understanding of a tragic issue that affects thousands of people around the world" (Bundy 2016).

Users can watch the experience on multiple devices, and it turned out to be an episodic mini-series in which the user turns into an investigative journalist visiting the protagonist at home, learning more about her experiences by her telling the user and then driving through town.

Fader: A Journalistic VR Platform

The core idea of Fader is to offer a journalistic production tool that is easy to learn so that journalists, storytellers and other creators can focus on the narrative of their stories instead of having to learn new software. Fader is a web application for creating and hosting 360° experiences. That means users can publish 360° stories as well as distribute these stories, for example via sharing on social networks, embedding on their website or by sending a link to the audience. Users can upload various media elements, such as 360° photos and videos, audio tracks, as well as legacy 2D media. These media items can be arranged in spherical scenes and users can add text layers and interactive hotspots to the story. Once they finish arranging their scenes, they can publish their Fader story. Because the story is a web-based project, users can go back and edit their stories or even add more scenes as the story is evolving.

To be able to render 360° and VR in a web browser, Fader employs A-Frame (A-Frame, n.d.) as its WebVR framework. A-Frame is a declarative wrapper for the open source WebGL library Three.js (Three.js, n.d.). Using a modern web browser, users can view VR content on their PCs or mobile smartphones using the respective hardware capabilities, such

as an HMD (Head Mounted Display) or a smartphone's tracking devices (gyroscope and accelerometer).

While it is feasible for developers or very tech-savvy journalists to build custom experiences with these technologies, it is still a significant technology barrier for most players in news and media organisations. Fader was developed to mitigate that barrier by providing solutions in the following areas.

Media Management

In order to display media in WebGL scenes correctly, a simple link to a resource file or an upload does not always suffice, as media elements must meet certain requirements. To be able to stream video content in VR, Fader applies a transcoding algorithm to all uploaded video files, while still preserving their original aspect ratio and size. Images which take up a significant portion of a system's memory are also transcoded in several resolution qualities to ensure they can be loaded and displayed effectively on devices with low bandwidth or performance specifications. The player logic of Fader makes sure that the corresponding media is displayed and any loading time upfront is reduced to a bare minimum.

Editor

To create and arrange content, Fader provides a WYSIWYG (What You See Is What You Get) editing interface called the Fader Editor.

It is the central hub for media upload, scene management and media library. Users can create a story and start out with an empty scene. Instances of media types can then be created and added to this scene. These media types can be 360 images or videos, flat images or videos, audio files, text, or interactive hotspots. A story can be composed of multiple scenes that play in the order they are queued. An exception from this order can be made if interactive hotspots are employed that allow direct links to other scenes.

Discovery and Publication

Each story can be made accessible via links and social media sharing or by making it discoverable on Fader's public story listing (Fader, n.d.-a).

Furthermore, stories can be embedded on the web pages and blogs of news and media organisations. Internal analytics enable users to measure the impact of their stories and track user behaviour.

There already is a great list of contributors to Fader. For example, France-based international broadcaster Euronews has created stories on Fader about the German federal election, the French election and about Climate Change, each in multiple languages (Fader, n.d.-b).

Furthermore, Germany-based international broadcaster Deutsche Welle also created interactive Fader stories, reporting for example about the federal elections in Germany (Fader, n.d.-c).

Conclusion

These are early days for VR storytelling techniques. With regards to investigative journalism, it is already discussed if journalists should focus on Virtual Reality (VR) or take broader approaches like Augmented Reality (AR), Mixed Reality (MR), or Extended Reality (XR) into account. It is obviously up to practitioners in the journalistic field to further try out, fail, evaluate, learn and adapt in order to develop fulfilling use cases of immersive journalism that live up to the ever-changing expectations of an audience that is not suffering from a lack of information, data, stories, and content.

References

A-Frame. (n.d.). [Homepage]. Retrieved from https://aframe.io/.

Beckett, S. (2015, January 28). *'Forensic Holodeck' Recreates Crimes in 3D.* Retrieved December 12, 2017, from http://www.bbc.com/news/av/technology-30992099/forensic-holodeck-recreates-crimes-in-3d.

Bogle, A. (2017, February 21). *Virtual Reality Will Take 'Maybe a Decade' To Become Mainstream.* Retrieved December 12, 2017, from http://mashable.com/2017/02/20/australia-virtual-reality-penetration/#E_GX2gVRRSqq.

Bundy, T. (2015, February 14). *Jehovah's Witnesses Use 1st Amendment to Hide Child Sex Abuse Claims.* Retrieved December 12, 2017, from https://www.revealnews.org/article/jehovahs-witnesses-use-1st-amendment-to-hide-child-sex-abuse-claims/.

Bundy, T. (2016, November 29). *Disfellowshipped: A Virtual Reality Experience.* Retrieved December 12, 2017, from https://www.revealnews.org/article/disfellowshipped-a-virtual-reality-experience/.

Deutschland Radio Kultur. (2017, April). *Stasiverhöre: Manipulierte Geständnisse*. Retrieved December 12, 2017, from http://blogs.deutschlandradiokultur.de/stasiverhoer/.

Dolan, D., & Parets, M. (2016, January 14). *Redefining The Axiom of Story: The VR And 360 Video Complex*. Retrieved December 12, 2017, from https://techcrunch.com/2016/01/14/redefining-the-axiom-of-story-the-vr-and-360-video-complex/.

Fader. (n.d.-a). Discover. Retrieved from https://app.getfader.com/discover/.

Fader. (n.d.-b). [Projects Created by Euronews]. Retrieved from https://app.getfader.com/discover?user=newmedia%40euronews.com.

Fader. (n.d.-c). [Projects Created by Deutsche Welle]. Retrieved from https://app.getfader.com/discover?user=deutsche_welle.

Fusion. (2015, August 17). *Virtual-Reality Experiment Lets You Walk Around the Ferguson Shooting Scene*. Retrieved December 12, 2017, from https://splinternews.com/virtual-reality-experiment-lets-you-walk-around-the-fer-1793844331.

GONE GITMO. (2007, August 15). *Gone Gitmo*. Retrieved December 12, 2017, from http://gonegitmo.blogspot.de/search?updated-max=2007-08-17T14:04:00-07:00&max-results=7&start=35&by-date=false.

Insta360. (n.d.). [Homepage]. Retrieved from https://www.insta360.com/.

Maschio, T. (2017, July 28). *Storyliving: An Ethnographic Study of How Audiences Experience VR and What That Means for Journalists*. Retrieved December 12, 2017, from https://newslab.withgoogle.com/assets/docs/storyliving-a-study-of-vr-in-journalism.pdf.

McCartney, G. (2011, March 2). *A News Medium in Which You Are There*. Retrieved December 12, 2017, from https://news.usc.edu/31775/a-news-medium-in-which-you-are-there/.

New York Times. (2016, November 1). *Introducing The Daily 360 From The New York Times*. Retrieved December 12, 2017, from https://www.nytimes.com/2016/11/01/nytnow/the-daily-360-videos.html?_r=0.

Papagiannis, H. (2017). *Augmented Human: How Technology Is Shaping The New Reality*. Sebastopol, CA, USA: O'Reilly Media. Retrieved from https://www.augmentedhuman.co/.

Rath-Wiggins, L. (2016, February 25). *Writing for VR? Three Things to Remember!* Retrieved December 12, 2017, from https://getfader.com/writing-for-vr/.

Reveal News. (2017). *Reveal News*. Retrieved 2017, from https://www.reveal-news.org/.

Ricoh. (n.d.). [Website for Theta Product Line]. Retrieved from https://theta360.com/de/.

Rouse, R., & Barba, E. (2017). *Design for Emerging Media: How MR Designers Think About Storytelling, Process, and Defining the Field*. USA: Springer International Publishing AG.

Three.js. (n.d.). [Homepage]. Retrieved from https://threejs.org.

Check for updates

CHAPTER 11

Visual Storytelling: Show, Not Tell? Show AND Tell

Jens Radü

Show, not tell. That is the lesson practically every young TV journalist is taught at the beginning of his or her career. The moving image depends on action and emotions. In many cases, spoken commentary is nothing but a compromise if there are no corresponding images. I too was taught this principle, back when I was a trainee at Westdeutscher Rundfunk in Cologne. And, of course, I applied it to all the films I was involved in after that until I started working at SPIEGEL ONLINE in 2006 where I developed the multimedia department. As glorious as it sounds—and it was truly a lot of fun—it was really just a one-man show in the first few months. A lot of work. A lot of groundwork. A lot of technology. And, most of all, the realization that it is not about "show, not tell," but about "show AND tell!" That's because the possibilities of online journalism make every reporter a broadcaster, publisher and radio station. And all of it happens at the smartphone level. Readers have become accustomed to finding out about events in multidimensional form, so that texts, films, photos, graphics, and 360° videos merge into multimedia worlds of experience.

J. Radü (✉)
DER SPIEGEL, Hamburg, Germany

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_11

When we designed the new version of our magazine app at SPIEGEL in 2015, we wanted to take this development into account. From the very beginning, the digital version of *Der Spiegel* was more than just a PDF that readers could flip through. The stories were expanded and augmented with videos, photo galleries, and interactive graphics. But we wanted even more in the new version: an independent multimedia report every week, told with the tools of online journalists and perfected for tablets and smartphones. The "Visual Story" was born. It was to be the home of visually opulent topics, current multimedia reports, and playful approaches, or what we called "play, not show."

The Visual Story about Trump (Fig. 11.1) was published in summer 2016, before the election. It was a scenario in video and animation based on Donald Trump's promises he made in his campaign: Muslim-ban, the wall on the border to Mexico, his perception of women. For the first time, we added an interactive, searchable part in this Visual Story: The



Fig. 11.1 "Der Trump-O-Mat". The Visual Story about Trump

reader could browse through hundreds of Trump's quotes, promises, and predictions.

A year later, we were astonished recognizing, how much of this scenario has become true.

Readers' reactions have been encouraging. The Visual Stories have become established as a fixed component of SPIEGEL's digital version, are among the most frequently used stories in an issue, and have even won a few prizes. But multimedia stories week after week, produced practically in assembly-line fashion? Can it work? Is it affordable? And what can we learn from it?

Be a Coder

Admittedly, this has led to some strange conversations. The previous week, I spoke by phone with our editorial coder, who had developed a tool for our team that enables us to produce many of our Visual Stories. This was our dialogue: "The autoslide is moving too fast in the section with the fragment. Can I modulate it in the last image?" "Yes. Set the autoslide time to 10,000, which should be sufficient for the audio." Got it?

Ever since our team began producing Visual Stories, a third category has crept into all the conventional journalistic stuff that (1) makes this job so endearing—research, interviews, filming, telling stories, and that (2) we accept as a necessary evil of the job—long and even very long conferences, forms, Windows updates, and I am not sure if it should belong to category 1 or 2: html. Of course, the programming language itself is nothing new, but it has experienced a well-earned renaissance in recent years, because it ensures that websites, articles and multimedia specials like our Visual Stories behave in similar ways, function and are not broken on all digital devices—smartphones, tablets, and laptops.

Our coder calls this *responsiveness*. However, it only works in combination with the right CSS, a sort of digital style guide. You have probably already noticed that I am getting into the weeds here. But in all seriousness, it is a lot of fun. Of course, the creativity of programming is somewhat limited, since the tool itself already exists as a basic framework. One section follows the next, and the user can proceed horizontally or vertically with his finger or the mouse. But we control what happens in the individual sections (text, video, photos, animation, 360-degree photos, and graphics) and what they look like (typography, colour, background,

and effects). This does not happen in a prepackaged editor that does all the programming work for you, but through code. "Under the hood," so to speak. And this becomes noticeable in a very practical way. If you often fall into a run-of-the-mill pattern while working with one of these prepackaged editors, our way requires direct access to the carousel of ideas. If something does not work, it does not happen. After all, everything is adjustable. And it can all be done without forms or constantly going back to IT.

This may all be simple stuff for the seasoned programmer. And I do not claim that I will be writing the journalistic killer app for Google's Playstore any time soon. But a basic understanding of the process of programming, the most important html terms and constantly working with them have enriched my job—and have saved time. I suddenly understand programmers. And, perhaps most importantly, our Visual Stories are getting better—from one line of code to the next, and from story to story.

Be a Dramaturg

Every journalist is Aristotle, if only unconsciously. But the dramatic theory that Aristotle devised under the Greek sun, the principle of the three-act structure—exposition, rising action with a climax and, finally, the catastrophe or resolution—is the basis of virtually all journalistic narrative forms. It is a tour de force. After all, it is the way Hollywood films, successful novels, and short stories work. And multimedia stories also work if they follow this pattern. In the beginning, we experimented with photocentric narratives for our Visual Stories. A photographer had brought along a wonderful report about his travels, and we built a story from the images, and then the photographer provided the commentary. A photo gallery with audio commentary. Ok. But there was a high reader exit rate. We saw the same high exit rate when we tried running nonlinear stories. The most effective way to carry the reader through the story was to apply Aristotle's rules for drama: introduction, build tension, (happy) end. Predictable? Yes. But it was an important lesson.

This is how the digital version of *Der Spiegel* looks like (Fig. 11.2). The table of content is the most popular way for our readers to get an overview of the magazine. While every article shown here is part of the printed edition as well, the Visual Story is digital-only content. Our readers like it: It normally ranks within the top ten of the most-read stories.



Fig. 11.2 The digital version of *Der Spiegel*

Be Effective (Use Editorial Planning Tools)

Where is the next Snow Fall idea lurking, and in which team or department is a project taking shape that merits pulling out all the stops? In modern editorial offices, this can no longer be negotiated on demand or in conferences. We need a planning tool for central topic management. Whether you use Trello, Slack or Desknet—a list of projects that are in process and status management (In which phase is the project? What is the next step? Who is involved?) facilitate the work of the entire editorial staff, but especially that of multimedia journalists, who typically work in multiple areas at the same time. This does not mean that technology replaces the face-to-face concept meeting or briefing. On the contrary, when applied consistently, technology actually frees up the space for such conversations, because it eliminates many general and relatively unproductive coordination meetings. Yes, the point is to talk less. But even more importantly, the point is to say more.

Be a Marketer

Multimedia storytelling is expensive. This variety of journalism requires films, photos, and graphics. In addition, the worthwhile stories that can be told with films, photos, and graphics do not take place exclusively in Germany. As a result, the costs of travel, equipment, and production often reach a level that exceeds the budget for purely written reportage several times over. After all, everything has to look good in the end and must function on all digital devices. The only way to ensure that multimedia storytelling still remains affordable is to use networks. We try to present our Visual Stories on all platforms and in all existing media and channels. Of course, our first target audience consists of the users and readers who pay for them, such as subscribers to the digital edition of SPIEGEL. But then we have the SPIEGEL ONLINE users who only access the free content, which, of course, includes trailers for Facebook, Gifs or Twitter cards. In this way, we can expand our reach and reduce costs through synergies. It is not just a question of presenting content across editorial platforms, but also a matter of production: video journalists at SPIEGEL ONLINE who accompany SPIEGEL authors, photographers who also shoot on research trips, and motion designers who are also thinking about web and TV. The central idea is to make outstanding research involving collaborative teamwork shine on as many channels as possible.

Be Tough on Yourself (Shorten)

Do 100% of the users of a multimedia story watch 100% of all the content offered? Every video, every photo gallery and all the interactive graphics? Of course not. The insight that emerges from user statistics may be a bitter pill to swallow for multimedia journalists with a sense of mission, and who may have invested weeks or even months in a project: In addition to time constraints, there are limits to most users' willingness to devote half an hour or even an entire hour of their full attention to a multimedia story. This was a learning process for us. When we design Visual Stories, we now structure them for 10 to no more than 15 min of user time. Even that is still a lot. But by alternating between various media, which—at least this is the intention—are almost imperceptibly intertwined, we hope to create a draw that holds the user's interest. The most important aspect of this form of designing stories is to keep things

short: videos no longer than a minute; text segments that fit on a smartphone screen. Do not overload the user with interactive graphics, and keep them simple. And make sure the whole thing doesn't last more than 15 min.

There are of course exceptions. But to be honest, users who are particularly interested in a topic will stick with it. The majority, however, lose interest in the long haul, no matter how finely chiselled the dramaturgy is.

An example for being tough to yourself: When our reporter came back from Italy with dozens of hours of interviews, video-footage and ideas, it was a process of weeks to cut the story about the steel factory in Taranto (Fig. 11.3) down so that it was (a) deep enough to match our journalistic demands, (b) exciting enough to hold our readers, (c) short enough to fit to our format: not more than 15 min. In the end, it all worked out—it was one of the most-read stories in this digital issue.

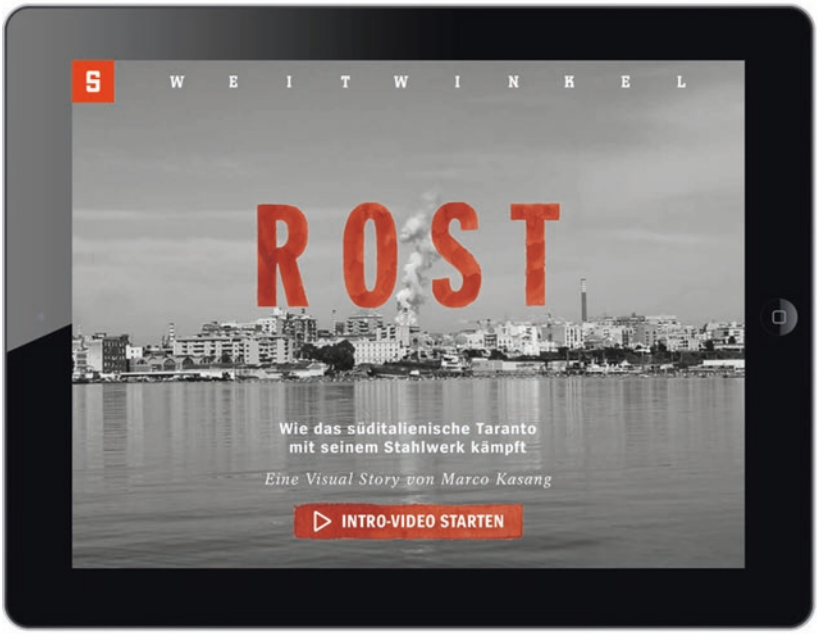


Fig. 11.3 "Rost". The Visual Story about an Italian steel factory

Be Surprising

When "Snow Fall" was published online in December 2012, multimedia journalism experienced its second bloom. In the early days of online journalism, the new technical possibilities, like audio slideshows, Gifs, and short animations inspired the imagination of the first generation of web journalists. Flash as the technical standard made new, non-linear narrative forms possible. But it was only "Snow Fall," the story of an avalanche accident in the Cascade Mountains of the United States, that provided the genre with its Multimedia-breakthrough. Not that everything was perfect. For instance, the combination of (very long) text, videos and interactive graphics is not optimally coordinated from today's perspective. And the story is not nearly as gripping and well-structured on the smartphone. Still, "Snow Fall" was a surprise. *The New York Times* had cleared its compartmentalized website for the story, so that there was no flickering advertising anywhere to be seen. Instead, there was a large-format, modestly animated teaser image, effectively typeset text, and an overwhelming 3D flight over the scene of the incident.

The form was new, the narrative logic was new, and the idea was new: longform journalism, combined with the tools of multimedia journalism. The surprise worked, leading to a Pulitzer Prize, millions of downloads and the establishment of an entirely new genre dubbed "scrollytelling." Five years later, the consensus is that the strength of "Snow Fall" was not just the quality of the story, but also the momentum: "Oh, this is new. This is interesting. This is cool." And for many journalists around the globe, the reaction was: "I want that too." But the goal should not be to copy "Snow Fall." Instead, it should be to surprise people and to be a laboratory for new storytelling ideas. From 360-degree to VR worlds of experience to story bots, the reader likes what he knows. But he also likes new things.

Be Hilarious

There will always be difficult topics: the mining of mica, the mineral that provides the sparkle in car paint and lipstick, in India by children in mines with few or no safety precautions. The extremely high suicide rate among adolescents in Greenland. Domestic violence against women in Russia. Subjects like these (the examples cited are Visual Stories from

the last two months) offer a lot of latitude for multimedia storytelling: highly emotional material, allowing for moving interviews and scenes, together with the escalation of exploitation, violence, and death. And, of course, audiences are not indifferent to stories like these. The Visual Story about domestic violence in Russia was in fourth place among the Top Ten most-read stories in the SPIEGEL digital edition. This is not without effect. Photographers who research available projects, multimedia journalists who apply for grants, as well as our topic conferences for the Visual Stories are focused on such serious and often dramatic topics. The idea is that if we are already pulling out all the stops for a big multimedia story, the topic itself should be as big and momentous as possible. Ironically, one of the most successful Visual Stories we have ever done was fairly light, even humorous. The British photographer Jocelyn Bain Hogg, who has documented the often eccentric lives of his fellow Britons in his images for decades, commented on a selection of his best photos in *a Visual Story* titled "Oh Britannia!" The subjects ranged from the crude tradition of the British hunting party to the drinking binges of elite students. The occasion, of course, was the Brexit vote. What exactly characterizes a British person? Bain-Hogg explores the depths of the British soul, both visually in the story and in his audio commentary (and he just happened to be drinking a gin-and-tonic in the process). He was surprised that the British ended up voting for the Brexit, and yet, as a Briton, he remained unperturbed: Keep calm and drink on.

More light topics, more humour, more courage to be funny, and to abandon the safe level of big, heavy topics—it is something we usually have to force ourselves to do. But it is worth it.

Be Your Biggest Fan

The day will come when you ask yourself the question of what it all means. Perhaps at three in the morning, when you realize that you cannot make your deadline. Or in a meeting with the agency, which unfortunately has completely misunderstood the briefing. Or at the moment, shortly before delivery, when the multimedia story your team has been working on for months suddenly no longer works because Apple has just issued a browser update. At moments like these, the best thing to do is to take a deep breath, place things in perspective (is it truly the end of the world?), analyse (how can we solve this problem?) and then carry on. Setbacks are part of everyday life for multimedia journalists. The story

has been wonderfully researched, the videos are done incredibly well, and the programming is clean—but then users do not understand the navigation. Or the link to share the story on Twitter does not work the way it should, and the viral hit turns into a non-starter. Failure is the big, admittedly ugly brother of progress. There is no innovation without mistakes. And in a journalistic genre that depends more on experimentation than any other, the offensive treatment of failures is part of the game. When we were about to launch the first Visual Story—a wonderfully animated background analysis of the oil crisis at the time—in December 2015 to mark the debut of the new *Der Spiegel* app, we were inundated with problems. The full-screen, animated teaser for the story could not be integrated into the app, the volume of data caused some test devices to crash, and then we had to work the current development of the oil price into the Visual Story. It was a 3 a.m. moment. Moments like that are easier to endure if you at least believe in yourself and the project. Everything works out in the end. And if it does not, you have not reached the end yet.

And now? We have just reinvented ourselves. The multimedia team at SPIEGEL has become a sort of visual lab of the sort that has been existing for some time at *The Guardian* and *The New York Times*, where infographics specialists, motion designers, photographers, coders, and VJs work together in an integrated team. The goal is to produce high-quality Visual Stories, interactive graphics, and animations for the entire company. Eye-catching, sometimes funny, and surprising. But most of all: simply good stories.

Check for updates

CHAPTER 12

Interactivity to the Rescue

Johanna Fulda

Online news has been around for a while now and we have seen its potential and its dangers. At first, the end of news was predicted, then the possibilities offered by digital technologies were cheered. Websites started to blink and scream, and the interactive graphic was born. We saw flagship websites guide us towards scrollytelling and then all of us wanted to tell stories the "New York Times way"—deep, investigative, full of media and wonder. But now that the novelty has worn off, it is time to start questioning interactive graphics. Do they actually achieve the prescribed reader engagement and pay off for the extra time and effort it takes to create them? Do they encourage loyalty towards the website and make readers tell all their friends about it? First and foremost: Do readers actually interact with graphics?

And here comes the shocking news: No, they do not. In fact, "nobody is interacting with news graphics" (Aisch 2016, 5:30).

When measuring the interaction rates on The New York Times (NYT) graphics, it turns out that only about 10–15% of readers click the buttons the creators intend to be clicked—no matter how salient they are or how many arrows point towards them. Reality is, as Archie Tse, Deputy Graphics Director at NYT, put it in his 2016 edition of "3 rules for

J. Fulda (✉)
Data Visualisation Design and Development, Vancouver, BC, Canada

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_12

visual storytelling" at Malofiej, Infographic World Summit: "If you make a tooltip or rollover assume that no one will ever see it" (Tse 2016). You cannot expect the reader to interact with your graphic. All critical information must be visible in the very first layer and not hidden behind interactivity.

At first, this might sound like sad news for information visualisation designers and journalists who are excited about the possibilities of interactivity in data visualisations. It allows for deeper stories, more context, and, when compared to static print graphics where you often have to drastically reduce the amount of data for the sake of formatting, a better medium for telling the full story. Interactive media allows us to show readers the full spectrum of information with many different viewpoints and perspectives, allowing the reader to go on their very own discovery trip—so why are they not clicking?

What Science Says

First let us look at how we define interactivity in this context and what studies have found about how interactive graphics can benefit engagement, recall, and memorability.

Definition: Interactivity

Interactivity is often defined as communication between sender and receiver on the Internet. This "revolutionary new medium" (Hoffman and Novak 1996)—the Internet—has transformed publishing from a strictly centralised model with powerful gatekeepers (the publishers) to a peer-to-peer based publishing model, where anyone can disseminate information or participate in conversations far beyond what the old "letter to the editor" allowed for.

In the early stages of online journalism Deuze (1999) already foresaw a cultural change caused by interactivity as "the main discerning characteristic of the online environment" (p. 378)—even though he was primarily talking about the new interactive possibilities for quicker communication between the journalist and its audience. He elaborates on this definition in his later work to three different types: "**navigational interactivity**", where the user has an influence on how they navigate the content; "**functional interactivity**", where the user can interact with other users or even the producer; and "**adaptive interactivity**" where the surfing behaviour is analysed to personalise the content according

to the users' preferences (Deuze 2003). Now a decade later we can add more interactive features such as sharing on social media (Hermida 2011), creating lists of most read or recommended stories to deal with the "limitless volume of information" (Singer 2014), polls, or the ubiquitous comment sections (Stroud et al. 2016).

We are focusing on the "navigational interactivity" here. Common ways in which a reader can interact with a visualisation are:

- **Inspect**: get details about something specific, for example, hover or click on an element to get more information
- **Connect**: show related items, for example, by clicking on one element to highlight similar ones for comparison
- **Select**: highlight element to keep track of it, for example, in an animated graph
- **Filter**: show something conditionally, for example, select one country from a drop-down menu
- **Abstract/elaborate**: show more or less detail by zooming in or out
- **Explore**: input a query to see "something else"
- **Narrate**: for example, a stepper-button that guides the reader to next part of the story (Boy et al. 2015, p. 1451; Yi et al. 2007, p. 1226)

Studies

There are several disciplines that care about interactive data visualisation. Many of these are within Computer Science, namely Information Visualisation (InfoVis) and Human Computer Interaction (HCI), which includes the technical implementation in addition to the visual encoding, understandability, and usability. Additionally, the digital humanities, communication studies, and journalism schools frequently investigate their usage and benefits as well.

There are a few studies that in particular look at whether there is a difference in engagement between static and interactive graphics. The results, while not earth-shattering, give some ideas:

- The Internet skill level plays a role in the use of interactive features (Chung 2008)
- When comparing an "interactive infographic, a static infographic and a table condition" small advantages were found for the interactive condition and disadvantages for the table (Milatz 2013)

- Interactivity (filter buttons in this case) can have a positive effect on "reducing political misperceptions". It can help in understanding a topic more accurately and inspire more thoughts (Geidner et al. 2015)
- Interactivity could lead to "more favourable attitudes towards an article for users low in involvement". For users that were engaged already it did not seem to make much of a difference (Wojdynski 2015)
- Self-paced animation control, where the user can decide when to start or stop an animation—which is a low form of interactivity—can reduce cognitive overload (Zhao and Huang 2013)

On the flip side, a 2010 study found that "interactive information graphics tend to overwhelm users with too much information and disregard well-known principles and rules of the old media and web design" (Burmester et al. 2010, p. 361). Further, unlike typographical style standards, such as The Chicago Manual of Style or the AP Stylebook, guidelines for interactive elements are inconsistent and vary widely across individual newsrooms (Günther and Scharkow 2014; Stroud et al. 2016).

In 2016, Dada tested some of those assumptions about how interactivity can improve graphics and online news and found "engagement benefits are restricted to participants who make use of interaction possibilities—and these are not in the majority" (Dada 2016, p. 52). She also finds in her experiments that usability and skills and motivations were responsible for more engagement and information recall.

What can we learn from this? It is advisable to know about the principles of user-based design, and in addition to maintaining the quality and meaningfulness of the content, one should follow the current design and usability standards and ensure that users know how to interact with the content. This is often not the case. Boy et al. studied whether users are naturally inclined to interact with visualisations and found that since the majority of users were unaware of the interactivity of the charts, they rarely used it (Boy et al. 2016).

What Practitioners Say

There is a lot of collaboration happening between Journalism schools, research labs, and news outlets. Still, it can be hard for practitioners to keep track of what science has found, and what new tools the research labs have developed, and for scientists to always see what problems

practitioners are facing. So, let us have a look at the other side, from the perspective of those who practice data visualisation every day in newsrooms and as reporters.

Interactives in the News

Data heavy reporting comes in many flavours, with terms like computational journalism, computer-assisted reporting, data-driven journalism, and digital investigative journalism (Coddington 2015). Despite the nuances distinguishing them, they are all very much intertwined with data visualisation. Mirko Lorenz describes the workflow for such journalism as:

- **Data Acquisition**: Scrape it, cleanse it, structure it, and start digging into it
- **Data Filtering**: Mine it for nuggets of information
- **Visualisation**: Create charts, graphics, or any kind of multimedia
- **Storytelling**: Tell the story around it, to vamp up the dry statistics and point out the interesting parts (Lorenz 2010).

The further along you move in that process, the more value you create for the reader since the raw data is transformed into something that is meaningful to the readers.

If you are using interactive data visualisations in the right way, they can turn an objective, impersonal, fact-based article into something much more personal and bi-directional. It could be "as if you were having a tête-à-tête with an expert on the data, patient enough to explain you everything" (Baur 2017).

Scott McCloud, the famous cartoonist and comics theorist, has many smart things to say about how to tell stories visually. The visualisation creator's job is to filter the information. In moving pictures, we use the time to direct the viewer, in comics we use space. "In interaction we are able to do it more explicitly, we're able to say hover your mouse over here or select this option and then you can see just this one trendline"—the one that is most interesting to you (Bertini and Stefaner 2017, 7:19).

Some topics are more interesting for this kind of storytelling than others. Murray Dick (2014) interviewed some interactives creators about it: Bella Hurrel, editorial lead at the BBC Specials team, says those topics are "'Big data' stories, concerning stories buried in large data sets… non-statistical process visualisations that help the audience understand

an issue… and personalised information generators, such as calculators" (p. 499). Also "outliers often make for good stories, as do stories which involve processes or personalisable information (such as school league tables)" (p. 500).

Emily Cadman from the Financial Times interactives team says that it is about "interactives that allow the user to explore something that affects them directly, or that encourage the exploration of complicated major topics" (p. 499).

Murray Dick concludes that whether interactives are used is less dependent on the subject but "due in part to the availability of 'newsworthy' data, and in part to newsroom dynamics". And since producing those interactives takes a lot of resources, "'important' stories are more likely to be selected for coverage" (p. 504).

Does Digital Journalism Need Interactives?

Visualisation is one essential part of this kind of digital journalism and very often those visualisations are interactive ones. In a content analysis about Canadian data journalistic projects we found that almost all (25 out of 26) projects were using some kind of interactivity in their stories (Young et al. 2017). In general, the tenor seems to be that interactivity is a core characteristic of digital journalism. Some of it is handcrafted, some of it uses freely available tools, such as Tableau, Datawrapper, or Google Maps (in fact Google Maps was by far the most used interactive technique in our Canadian content analysis). Since not all newsrooms have dozens of people in an interactive lab, they often rely on the eagerness of individuals willing to invest extra time in side projects, and in learning to work with available tools on their own (Hermida and Young 2017). Unfortunately, this can potentially lead to the journalistic output "to be shaped less by what could be considered the best way of representing/exploring the data and more by what can be done and is available for free" (Young et al. 2017, p. 13). The study indicated one or both of two things for newsrooms in Canada (and most likely over its borders)

1. There is a lack of knowledge about the principles of information visualisation, of how to effectively represent data and/or
2. There is a lack of time/people/experience to create customised solutions, so that freely available tools and platforms become shaping factors

Was there to be a better understanding of the principles and more consideration going into the decision whether to use interactives at all, we might see less unjustified interactives and more "good-old" static visualisations, that deliver the message just as nicely (see also next part "Rules of thumb").

There already seems to be a trend away from interactives, and towards creating expressive static graphics, where no information is "hidden" behind tooltips. Priya Krishnakumar, a visual/data journalist at the LA Times, says "lately I'm trying to hew more towards static if I can" (Schwabish 2017, 7:40). Martin Stabe, the head of interactive news at The Financial Times says that they too are trying "to avoid using interactivity where possible because it is more important to make something responsive than adding that extra layer of complexity (Luyckx 2017).

Complex Visual Storytelling Needs Interactivity

Data Visualisation is there to clarify—not to simplify. Alberto Cairo's aims to get that misconception straight and says that while it is necessary to remove unnecessary information and "simplify" the message, we must be extremely careful in ensuring that nothing essential is removed. Sometimes, clarity can require additional data to provide the appropriate context—a task extremely well suited for interactivity (Rogers et al. 2017).

How does this argument hold up to the fact that only very few people will get to that additional information—because, remember? Nobody clicks!

Simon Rogers (Guardian Datablog creator, now at Google) argues that the number of people using interactive visualisations (10–15%), is not much lower than the number of people reading the full article and therefore, this should not be the main argument against using interactives. The small percentage of readers that actually click the buttons are the ones you are creating it for—the interested ones (Aisch 2017; Rogers et al. 2017).

Clearly, not every topic lends itself to this kind of time investment. In many cases, readers visit news websites on their mobile phones to catch up with the latest and hottest while riding the subway to work. But the more complex, interesting, and data-heavy the topic, the more it does lend itself to a deeper interactive investigation. Matt Daniels, member of Polygraph (a visual storytelling incubator) and founder of The Pudding

(a journal for visual essays), says "I want the thing to be so deep, and complex, and interesting that it is like a Lord of the Rings moment" (Rogers et al. 2017, 24:30)—which is when you want to be alone with your screen, getting fully emerged, digging in, understanding everything.

Rules of Thumb

Interactivity can greatly enhance data journalistic pieces. While there is no need to despair just because not everybody engages with it, we should preach caution in the way it is used. The decision to use interactivity has to be deliberate, conscious, and considered. A few rules of thumb might help make the call.

Justify Interactivity

"Interactivity is an extra layer for visualisations, that will make [a data visualisation] more complex and therefore more time-consuming" says Martin Stabe from the Financial Times (Luyckx 2017). Think about what the benefits of an interactive would be. In many cases, a well-curated static graphic is way more useful and accessible than a blown-up interactive.

Bottom line question: Does this particular piece really need to show more data than we can fit in a static graphic?

No Arbitrary Interactivity

The availability of easy-to-use data visualisation tools has led to their widespread use, often by relatively novice users who stick to the default settings. These settings are sometimes distracting: for instance, while Google Maps understandably allows zooming and panning, sometimes we want to show an explicit part of the map and prevent the user from getting lost by accidentally scrolling over it. So, in this case a conscious decision would be to disable the default zooming and panning and only show the one area of interest.

Other chart libraries that enable the user to quickly create bar/line/scatter/pie charts advertise themselves as having "easy interactivity". This is great, but also really unnecessary at times. Simple pie charts with three pieces of clearly visible static information that change colour and are highlighted when the user hovers over them—interactivity overkill for no discernible extra information.

One might say: Well, it does not hurt, does it? But it is a misuse of interactivity. Actually, it can hurt comprehension by forcing the reader to search for meaning behind the interaction.

Every visual decision generates meaning. "Our eye assigns meaning to every detail that is there. […] We assign significance to it at least subliminally. We're going to think that that matters, that we are being told something about that little nugget of data", says Scott McCloud (Bertini and Stefaner 2017, 30:30). So, be careful not to generate the wrong cues.

Bottom line: Disable interactivity if it does not have a purpose.

Lean Forward or Lean Back

Within interactive data visualisations the degree of interactivity can vary a lot. There are "interactive" pieces that only have one "narrate" button, for example, to switch a visualisation from the current to the previous year. Segel and Heer call this an "author-driven" approach (2010, p. 1145), as it uses linear storytelling and the author guides the reader through the visualisations. In contrast to that there's the "reader-driven" approach, where the reader has many more options to interact and explore.

Some topics lend themselves better to the reader-driven method than others; for example, stories where the reader expects to get information quickly without having the time for exploration are a poor fit for this model. Scott McCloud talks about a conflict between driving and being driven. When considering the options of Virtual Reality, he says "Either I'm in control, or the author is. Either I lean forward or I lean back. I think that we tend to be a little uncomfortable when it's flickering somewhere in the middle" (datastories, 28:10). Similarly, the BBC Specials team found that when "taking a prescriptive narrative approach, more readers are likely to complete all stages of the interactive—so this approach is often preferred to experimental interactive graphics" (Dick 2014, p. 502).

Bottom line: Make the decision who is driving: author or reader.

Bart and Lisa Approach

In information design it is a good idea to create something that works on different comprehension levels. Former New York Times graphics editor Jennifer Daniel explains an idea that they called the "Bart and

Lisa approach" (in a reference to "The Simpsons" TV series). A graphic should be attractive to Bart—with his rather low-attention span—as well as to Lisa, who is interested in details and wants to understand the context and learn something new. So for interactive graphics: Bart should be able to get the message without having to move the mouse at all; whereas Lisa can dig way deeper to satisfy her curiosity (Miller 2016).

Bottom line: The important parts have to be there even without interacting.

Have a Fallback Solution

The greatest interactive graphics sometimes use the latest web technologies, need the biggest space, or load a huge amount of data. Unfortunately, we cannot assume that the readers always have the latest browser, the biggest screens, or the fastest internet. So if we decide to go for it anyway, there has to be a fallback solution.

This becomes harder the more complex the data and the visualisation is, and it is likely that information gets lost along the way (Aisch 2016, 10:30). But rather than showing a message like "Please go to your desktop computer to enjoy the full experience", or "Sorry but this only works on the latest version of Browser XY", it would be good to have a static fallback image that shows as much as possible.

Often this will be the case for mobile devices because elaborate interactives tend not to work as well on them. Since mobile is becoming the number one way how news is consumed, this brings up a whole new abundance of challenges, such as screen size, internet speed, the fat finger problem, and an (even more) limited attention span. This is beyond the scope of this chapter, but some interesting resources in this space are: Julia Smith's "Data Viz for All: Best practices for mobile, accessible interactives", and more great examples, patterns, and best practices can be found on the MobileVis website.

Bottom line: Consider all the platforms and devices and offer at least a fallback solution.

Suggest Interactivity

As we have seen earlier, interactive features often remain undiscovered by the audience (Boy et al. 2016). If you are deciding to use an interactive graphic, make it clear how this interactivity is supposed to work. Alastair Dant, lead interactive technologist at the Guardian, mentions

the importance of providing "a clear journey through the data or through whatever is in your interactive experience" (McAthy 2012). Consequently, we should include annotations, instructions, directions, or suggest the interactivity with visual means—Boy et al. formulated an extensive table with visual cues that designers can use to suggest interactivity to users.

Martin Jefferies, the chief reporter at the KM Group, concurs: "You need to tell them what they've got to click on" (McAthy 2012).

Bottom line: Do not assume the reader will intuitively know what to do

Conclusion

Interactives are a great way to present complex stories to the discerning reader. As with every new technology, however, they have often been overused. Designers sometimes opt for form over function and spend too little time thinking about usability and the principles of data visualisation, which can lead to interactivity being hidden or unnecessary. This chapter aims not to diminish the value of interactives, but to highlight some of the pitfalls and real costs associated with it, and suggests some rules of thumb for practitioners who are uncertain about whether it is the right choice and what to consider when working with complex and data-heavy topics.

References

Aisch, G. (2016, September 13). *Information+ Conference: Gregor Aisch* [Video file]. Retrieved from https://vimeo.com/182590214.

Aisch, G. (2017, March 31). *In Defense of Interactive Graphics* [Web log post]. Retrieved from https://www.vis4.net/blog/posts/in-defense-of-interactive-graphics/.

Baur, D. (2017, March 13). *The Death of Interactive Infographics?* [Web log post]. Retrieved from https://medium.com/@dominikus/the-end-of-interactive-visualizations-52c585dcafcb.

Bertini, E., & Stefaner, M. (Producers). (2017, July 17). *Datastories 102: Understanding Comics and Visual Storytelling with Scott McCloud* [Audio podcast]. Retrieved from http://datastori.es/102-comics-and-visual-storytelling-with-scott-mccloud.

Boy, J., Detienne, F., & Fekete, J. D. (2015). Storytelling in Information Visualizations: Does It Engage Users to Explore Data? In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1449–1458). New York, NY: ACM.

Boy, J., Eveillard, L., Detienne, F., & Fekete, J. D. (2016). Suggested Interactivity: Seeking Perceived Affordances for Information Visualization. *IEEE Transactions on Visualization and Computer Graphics, 22*(1), 639–648. https://doi.org/10.1109/TVCG.2015.2467201.

Burmester, M., Mast, M., Tille, R., & Weber, W. (2010, July). How Users Perceive and Use Interactive Information Graphics: An Exploratory Study. In *Information Visualisation (IV), 2010 14th International Conference* (pp. 361–368). London: IEEE.

Chung, D. S. (2008). Interactive Features of Online Newspapers: Identifying Patterns and Predicting Use of Engaged Readers. *Journal of Computer-Mediated Communication, 13*(3), 658–679.

Coddington, M. (2015). Clarifying Journalism's Quantitative Turn: A Typology for Evaluating Data Journalism, Computational Journalism, and Computer-assisted Reporting. *Digital Journalism, 3*(3), 331–348. https://doi.org/10.1080/21670811.2014.976400.

Dada, J. (2016). *Engaging to Explore? An Online Experiment to Investigate the Impact of Interactivity in Data Visualizations.* Unpublished Master's thesis, University of Oxford.

Deuze, M. (1999). Journalism and the Web: An Analysis of Skills and Standards in an Online Environment. *International Communication Gazette, 61*(5), 373–390. https://doi.org/10.1177/0016549299061005002.

Deuze, M. (2003). The Web and Its Journalisms: Considering the Consequences of Different Types of Newsmedia Online. *New Media & Society, 5*(2), 203–230. https://doi.org/10.1177/1461444803005002004.

Dick, M. (2014). Interactive Infographics and News Values. *Digital Journalism, 2*(4), 490–506. https://doi.org/10.1080/21670811.2013.841368.

Geidner, N., Pjesivac, I., Imre, I., Coman, I., & Yuran, D. (2015). The Role of Interactive Graphics in Reducing Misperceptions in the Electorate. *Visual Communication Quarterly, 22*(3), 133–145. https://doi.org/10.1080/15551393.2015.1069195.

Günther, E., & Scharkow, M. (2014). Recycled Media: An Automated Evaluation of News Outlets in the Twenty-First Century. *Digital Journalism, 2*(4), 524–541.

Hermida, A. (2011). Mechanisms of Participation. In J. B. Singer (Ed.), *Participatory Journalism: Guarding Open Gates at Online Newspapers* (pp. 11–33). Malden, MA: Wiley.

Hermida, A., & Young, M. L. (2017). Finding the Data Unicorn: A Hierarchy of Hybridity in Data and Computational Journalism. *Digital Journalism, 5*(2), 159–176. https://doi.org/10.1080/21670811.2016.1162663.

Hoffman, D. L., & Novak, T. P. (1996). Marketing in Hypermedia Computer-mediated Environments: Conceptual Foundations. *Journal of Marketing, 60*(July), 50–68.

Lorenz, M. (2010, June 29). *Data-Driven Journalism: What Is There to Learn? (Stanford, June 2010)* [Talk Slides]. Retrieved from https://www.slideshare.net/mirkolorenz/datadriven-journalism-what-is-there-to-learn.

Luyckx, D. (2017, February 9). *3 Ways to Make Data Visualisations and Interactives Work on Mobile*. Retrieved from https://www.journalism.co.uk/skills/3-ways-to-make-data-visualisations-and-interactives-work-on-mobile/s7/a699519/.

McAthy, R. (2012, August 1). Visual Journalism: Advice on Building Interactives and Engaging the Audience. Retrieved from https://www.journalism.co.uk/news/tips-engage-the-audience-with-interactive-journalism/s2/a549988/.

Milatz, M. (2013). *Moving Graphics: The Effects of Interactive Infographics on Media Users' Recall Accuracy*. Unpublished master's thesis.

Miller, M. (2016, August 12). *The "Bart and Lisa" Theory of Information Design*. Retrieved from https://www.fastcodesign.com/3066148/the-bart-and-lisa-theory-of-information-design.

Rogers, S., Cairo, A., & Daniels, M. (2017, March 31). *News Lab Data Visualization Round Up: March 2017—With Special Guest, Matt Daniels* [Video file]. Retrieved from https://www.youtube.com/watch?v=UqLAxpumENs.

Schwabish, J. (Producer). (2017, June 6). *PolicyViz Episode #89: Priya Krishnakumar* [Audio podcast]. Retrieved from https://policyviz.com/podcast/episode-89-priya-krishnakumar/.

Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics, 16*(6), 1139–1148.

Singer, J. B. (2014). User-generated Visibility: Secondary Gatekeeping in a Shared Media Space. *New Media & Society, 16*(1), 55–73. https://doi.org/10.1177/1461444813477833.

Stroud, N. J., Scacco, J. M., & Curry, A. L. (2016). The Presence and Use of Interactive Features on News Websites. *Digital Journalism, 4*(3), 339–358. https://doi.org/10.1080/21670811.2015.104298.

Tse, A. (2016, March 21). *Why We Are Doing Fewer Interactives* [talk slides]. Retrieved from https://github.com/archietse/malofiej-2016/blob/master/tse-malofiej-2016-slides.pdf.

Wojdynski, B. W. (2015). Interactive Data Graphics and Information Processing. *Journal of Media Psychology, 27*(1), 11–21. https://doi.org/10.1027/1864-1105/a000127.

Yi, J. S., ah Kang, Y., & Stasko, J. (2007). Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics, 13*(6), 1224–1231. https://doi.org/10.1109/tvcg.2007.70515.

Young, M. L., Hermida, A., & Fulda, J. (2017). What Makes for Great Data Journalism? *Journalism Practice*. https://doi.org/10.1080/17512786.2016.1270171.

Zhao, L., & Huang, T. (2013). Experimental Research on the Effectiveness of Animated Instruction. In *3rd International Conference on Multimedia Technology (ICMT-13)* (pp. 562–569). Amsterdam: Atlantis Press.

Check for updates

CHAPTER 13

Visual Analysis: Verification via Geolocation and Photographs

Eliot Higgins

With the wealth of photographic and video content now being published the verification of this imagery has become increasingly important. The basic question of where an image was captured has become key to the process of verification, and through the technique of geolocation it is often possible to find the exact position an image was recorded.

Geolocation involves using all available material to confirm the location of an image. Often this involves satellite imagery available on a range of services, such as Bing Maps, Google Earth, Google Maps, and Terraserver. On the ground imagery, the location of which has already been established, is also frequently used, and examples of this include Panoramio and Google Street View.

However, the first challenge is, in a geographical sense, knowing where to start looking. Images will often come with additional information, for example, the town or city the image was taken in, or they

E. Higgins (✉)
Bellingcat, Leicester, UK

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_13

will be posted on a social media profile belonging to an individual in a specific location. What must first then be established is what the original source of the image is. While in some circumstances you may be aware of the original source, many images now shared online are often reposted without the original source included. This can be for a variety of reasons, but the need to establish the original source whenever possible is always there.

In the case of photographs, it can be extremely simple. Reverse image search functions on Google Image Search or TinEye, both freely available and easy to use, is the first tool to use in the search for the original source of images. Browsers such as Chrome now allow users to simply right-click an image and select reverse image search, so it could not be easier. These searches will bring up a list of results, and often it will be a simple matter of looking through those search results to establish whether or not an image is new, or has been previously posted online. One pitfall here is that dynamically generated pages can lead to misleading results. A YouTube page will list other videos related to the video currently being viewed, with a dynamically generated list that includes preview images which Google Image Search will find. This means that when the search engine has crawled that site, it has resulted in a dynamically generated page being made, and the image search results are based on when the search engine crawled that site, so it will include all the time that image has appeared on the page as a related video on the page of another video, and might not even appear at this point in time.

Another issue is that a page might be created on a date, but the content is added to it at later dates, for example on an internet forum thread. Often you will then see that the date given in the Google Image Search result is the date when the webpage was created or when it was first crawled by the search engine, not when the image was posted on that page. In the case of the downing of Flight MH17 in Ukraine, it was falsely claimed that one image showing a Buk missile launcher in Ukraine and linked to the downing of the aircraft was published the day before the attack, not on the day of the attack as many people had claimed, because the reverse image search showed a search result that listed the previous day. In fact, the page it was posted on had been created on a previous day, but the update with the image was made later.

Videos can be harder to find. Unlike photographs, there is currently no reverse search platform for videos, but there is at least one work-around. Amnesty International's YouTube Dataviewer will

extract the preview frames from a YouTube video, allowing them to be easily reversed image searched. This will find other versions of the video in question, by searching for the preview frames in those versions of those videos. However, there is a major limitation to this: It will not be accurate if the video is edited, shortened, or extended with additional footage.

If you are able to find the original source of the image, it then may be possible to find more information about it. The simplest information to find is whatever was posted along with the image. For example, an Instagram image or Facebook post might have the location listed in the description or discussed in the comments. The next step would be to look at the profile it was posted on. Does it usually post about one location? Is the person who posted it talking about their movements?

This may, of course, not even be necessary depending on what the image actually shows. A street sign or shop name might give a crucial clue to where the image was recorded. These can be easily searched for on Google along with the name of the locations it may be into see if there are any results, which will often point to specific locations. In some cases, Street View imagery may be available, at which point it should be possible to verify the location in the image by comparing it to the Street View imagery.

If you do not get lucky you then have to start thinking about how to narrow down the possible locations it could be in. If you have the general location that can be useful, but finding the precise location requires additional effort.

Begin first by identifying large features and distinct landmarks. During the conflict in Libya in 2011, a video was posted online by rebels claiming to show them entering the town of Tiji, previously controlled by Colonel Gaddafi's forces. Establishing the exact location it was filmed was possible, thanks to two major landmarks visible in the video. First, the video was filmed outside a large mosque with a dome and minaret, which would be clearly visible in satellite imagery. The major landmark was the road next to the mosque, two wide lanes, with a slight bend in the road clearly visible. By examining satellite imagery of Tiji on Google Earth it was immediately apparent that a major road ran through the centre of the town, and a closer look at the road was clear enough to show individual cars on the road, with each lane being three car widths wide, fitting with what was visible in the video. Also visible in the satellite imagery was a slight bend in the road, and close to that was a mosque, with a dome and minaret.

These two larger features had taken us to the same location the image was recorded, but next, the video was re-examined to identify other, smaller, features. The position of the curb, the location of utility poles (visible by the shadows they cast), the position of buildings and trees. While these features would not have led us to the location, they allowed us to confirm the location, and precise position of the camera.

Sometimes it takes more than satellite imagery to find a location. When investigating the downing of flight MH17 in Ukraine, a number of videos and photographs of a Buk missile launcher being transported through separatists' territory had been posted online. With various claims being made about what had happened, it was important to ensure all images were verified, especially as the images appeared to show the missile launcher on the way to the launch site from which it shot down MH17.

One image, a photograph taken from a garage forecourt, showed a Buk missile launcher being transported by a low-loader truck. The photograph was being shared on various social media platforms, with claims that it was either taken in the town of Torez, or in the town of Snizhne.

The first clue was a partially visible shop sign. Although the sign was obscured, there were only reasonably few possibilities of words it could be, and a Google search of those words and the town names resulted in only a few viable leads, one of which was a Wiki listing streets in Ukraine with the shops on them. This provided the street name, which was, along with the shop name and town name, searched for on Google, resulting in a new result. This was a court document detailing a fight that had taken place in the shop, giving the full address. Searching for the address and shop name led to a location visible on satellite imagery just across the road from a garage forecourt, the same one the photograph had been taken from.

In addition, during the Google searches two videos were discovered, filmed by a local on a dashboard camera from his car as he drove around the area, which he had uploaded to YouTube, listing the streets he drove down in the description of each video. Now in addition to the satellite imagery match, it was also possible to match imagery from the videos to the location in the photograph. Because the videos showed an entire journey through the area, it was possible to further verify where the video was filmed by matching off various landmarks along the route filmed on the dashboard camera, which in turn allowed the location of the photograph to be more certain.

Understanding how information is organised and shared online can also be useful. A bus visible in an image would suggest the location of



Fig. 13.1 Crowdsourced geolocation via Twitter and Google Street View

the image is on a bus route. In many countries information like bus routes, tramlines, etc. are available online, and often include maps. On Google Earth and Yandex Maps tram and bus routes are sometimes visible, which can dramatically reduce the number of areas needing to be looked at.

There is also less obvious information that has been organised online. During an ISIS social media campaign, supporters of ISIS in Europe were encouraged to take photographs of themselves holding a piece of paper with a pro-ISIS hashtag on it, and the town or city they were in. The location of one photograph in Paris was easily found because a Suzuki logo was partially visible, and even though it was partly obscured and out of focus, it was still easily identifiable. A Google search for "Suzuki Paris" brought up the locations of Suzuki dealerships in Paris, each of which was visible on Google Street View, and from there it was easily possible to match other features in the background of the photograph (Fig. 13.1).

In another photograph from the ISIS social media campaign a photograph from Münster, Germany, was located because of an advertising pole visible in the background. In Germany there is a website that lists



Fig. 13.2 Crowdsourced geolocation of advertising sites via Twitter

the locations of advertising sites, intended to help buyers find advertising positions across Germany. It was possible to search for all advertising poles in Münster, immediately giving a list of potential locations to search, each visible on Google Maps satellite imagery, which could then be systematically reviewed, until the location of the ISIS supporter who had taken the photograph was established (Fig. 13.2).

Ultimately, the only way to develop geolocation skills is simply just to do it. When I began verifying images, I found it extremely helpful to write out the steps I had taken to verify the image I was working on. This allowed me to really understand the process in my mind, and to develop my own approaches to assessing and verifying material. Geolocation is not a complicated discipline, so the best place to start is just picking an image or video, and figure out exactly how to prove where it was taken. Just going through the process, even with a simple geolocation, can build skill, and help you develop an eye for details in images you might not have noticed before. It also pays to revisit old cases, reviewing images you may have looked at months or years ago, as often you will spot new details and other paths to geolocating the image that you would not have seen before. Do this, and soon enough you will develop the skill set and knowledge required to be an expert geolocator.

Check for updates

CHAPTER 14

eyeWitness to Atrocities: Verifying Images with Metadata

Eleanor Farrow

This case study looks at how the eyeWitness to Atrocities project has collaborated with investigators to strengthen their reporting and collect court-admissible evidence using the eyeWitness app.[1] eyeWitness to Atrocities (eyeWitness) is a free smartphone camera app designed for individuals investigating human rights violations and international crimes. Developed under the auspices of the International Bar Association, the app automatically collects basic information required by courts to authenticate images. Metadata-enriched images taken with

[1] In this chapter "eyeWitness to Atrocities" refers to the free android smartphone camera app of that name. The eyeWitness to Atrocities project is also the name of the independent registered UK charity comprised of a small team of lawyers funded by the International Bar Association responsible for maintaining the app, collaborating with human rights investigators for effective use of the app, and use of data collected using the app for justice and accountability. The term "human rights investigators" is used in this chapter as an umbrella term to describe investigative reporters, nongovernmental organisation (NGO) journalists and human rights researchers documenting human rights violations and international crimes.

E. Farrow (✉)
eyeWitness to Atrocities, London, UK

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_14

the app can also support robust reports. The organisations that eyeWitness works with collect information using a combination of technology tools and traditional investigation methods. Despite the promise of new technologies for human rights investigations, their deployment comes with challenges (Land et al. 2012; Tuckwood 2014; Raymond and Sandvik 2017; Aronson 2017). This case study also includes reflections on lessons learned during the eyeWitness project and outlines some key considerations when adopting technology tools in human rights investigations.

Verification and Human Rights Reporting

Smartphones and social media platforms have increased the flow of information about human rights violations around the world. At the same time, international news bureaux have closed, and the number of professional foreign correspondents has reduced due to economic pressures (Constable 2007; Sambrook 2010). In addition to the lack of resources needed to retain foreign correspondents, there are difficulties in accessing certain locations (Reporters Without Borders 2013; Allan 2017; Koettl 2017). Non-governmental organisation (NGO) researchers often step into the breach to provide information (Powers 2015). News editors are increasingly reliant on user-generated content (UGC) and citizen reporters, particularly in areas of armed conflict (Platt 2014; Mast and Hanegreefs 2015).

One result of these trends is that media outlets are more dependent on digital images from unknown sources that are easily manipulated. News editors often have to judge whether to publish an image that they have not been able to verify. Errors can occur with potentially serious reputational or legal consequences (Shapiro et al. 2013, p. 664; Safran 2005). Yet, disclaimers that an image could not be verified potentially weaken a report. Media organisations are actively seeking UGC, and there is a flood of publicly available material posted on social media sites. This has increased the need for image verification.

The verification of facts and sources is fundamental to both investigative journalism and human rights fact-finding. This is particularly the case where official, organised, state-level denial of human rights violations and international crimes occurs (Cohen 2001, p. 10). News organisations and human rights organisations depend on their reputation for credibility (Land 2016, p. 409; McPherson 2015b, 2016). With worldwide concerns about “fake news” and propaganda, and many images

appearing on social media out of context or edited, checking the authenticity of images remains crucial (Fabijanić et al. 2016). Koettl describes this as "the 'lemon problem'—the risk of using misinformation that can discredit an entire research project" (Koettl 2017, p. 38).

Investigative journalists, in particular, have been said to adopt "a distinctly quasi-scientific approach, setting out to prove or disprove a hypothesis through triangulation of information from various sources" (Cribb et al. 2006; Hunter 2011, as cited in Shapiro et al. 2013, p. 659). Traditional methods of authentication include assessing the reliability of individual sources and cross-checking with two or more independent sources to corroborate facts (Golder and Reich 2017; Brandtzaeg et al. 2016; Shapiro et al. 2013; Hermida 2015; Wardle et al. 2014, p. 54). More time-intensive techniques available to verify images include expert analysis to detect edits, site visits to confirm the location, using satellite images, weather reports, shadows and computer modelling (Goldsmiths University "Forensic Architecture" 2011; Pantti and Sirén 2015, p. 496; Bell 2015).

Many newsrooms now have verification processes and resources in place, and it is hardly surprising that there are a growing range of image authentication projects and tools (Silverman 2012). Indeed, some have argued that "an image verification industry" now exists (Pantti and Sirén 2015, p. 498). Resources and projects addressing this issue include the Verification Handbook, Storyful, Verifeye Media, Checkdesk, FourandSix, Citizenside, Amnesty International Citizen Lab, First Draft News, Witness, StopFake and Bellingcat (Silverman 2012; Bell 2017; Allan 2017; Witness, n.d.).

A growing range of new technology tools aim to facilitate and streamline image verification in response to constraints on time and resources. Tools designed to verify images after they have been taken include Amnesty International's YouTube Data Viewer, Google reverse image search, TinEye, Findexif and FotoForensics. When images are shared via Facebook, Twitter, YouTube, or WhatsApp, key information about the source and context are often stripped or misattributed. Smartphone apps designed to collect key metadata information such as the time, date and location at the point of capture include eyeWitness to Atrocities, International Evidence Locker, Truepic, Whistler, the Guardian Project's Camera V and Proofmode. These tools are potentially useful in tackling this "disembodiment of information" (McPherson and Probert 2017, p. 267; McPherson 2015a, p. 21; 2015b, p. 198).

Legal "Reliability"

Human rights investigators often collect information with two purposes in mind. In the short term, information can be used in reports to bring issues to the attention of the public and key actors including "gatekeeper" human rights organisations, governments, embassies, international and intergovernmental NGOs (Carpenter 2010, p. 205; Bob 2008, p. 6).[2] In the long term, this data can also be put forward to seek social change, and if collected properly, to seek justice and accountability. For those collecting digital images for reporting on violations in the near term, adhering to more rigorous verification standards can help to increase their future impact (Lowe 2015).

The eyeWitness project focuses on justice for serious violations of international law; the design of the eyeWitness camera app is therefore founded on the core admissibility requirements for photo and video evidence in international, regional and national courts. In particular, the eyeWitness project reviewed the requirements of the International Criminal Court (ICC), the International Criminal Tribunal for the Former Yugoslavia (ICTY), the International Criminal Tribunal for Rwanda (ICTR), the Extraordinary Chambers of the Courts of Cambodia (ECCC), and the European Court of Human Rights (ECtHR).[3] These courts draw on both national common law and civil law traditions.

Even though the admissibility requirements of different courts vary, there are some shared principles. Firstly, the images must be relevant, or of probative value, meaning they help to prove or disprove a fact at

[2] See Carpenter (2010, p. 205). Carpenter describes gatekeepers as "any global actor densely connected to a particular issue network and thereby possessing particular influence over the issue agenda in that network" (Bob 2008, p. 6). Bob defines gatekeepers as "entities at the core of the human rights movement, whose support for a claim can boost it substantially. Typically these are organizations with the largest budgets, best staffs, and greatest credibility in the rights movement. Among them…major NGOs such as Amnesty International and Human Rights Watch; international organizations such as the UN's Office of the High Commissioner for Human Rights and other prominent international bodies; and human rights intellectuals".

[3] See Ashouri et al. (2014) and Freeman (2018), for research on the use of digital evidence in international criminal courts. See also the International Criminal Court e-Court Protocol from the Office of the Prosecutor Regulations ICC-BD/05-01-09 and Regulations of the Court ICC-BD/01-01-04 (26 May 2004) as amended by Prosecutor v. Callixte Mbarushimana, Case No. ICC-01/04-01/10, Decision amending the e-Court Protocol.

issue in the case. Secondly, the courts require the images to be "reliable". In assessing reliability, courts consider whether images are original and unedited from the source, the chain of custody record, and information about the time, date and location at the point of capture. The chain of custody means information to establish the integrity of the image by tracing where it has been, who has had access and how susceptible to tampering it has been until offered in evidence (Umberg and Warden 2014; The Engine Room 2016a, p. 36; Freeman 2018, p. 292). The eyeWitness camera app was designed to address these authentication requirements.

At the international level, authorship of images is considered important in determining the weight given to the evidence, but the absence of information about the author is not a bar to admissibility. There is a need to balance protecting witness identity and establishing the source (Ashouri et al. 2014, pp. 3–11). The eyeWitness app therefore provides optional anonymity, and the project team consults those using the app who chose to provide contact information before sharing their data with a court. However, those who provide contact details may also be called upon to testify in support of their images. Consultations with legal professionals have confirmed that images captured with the app would be more easily admissible. At the time of writing it is anticipated that information collected with the app will go to trial for the first time in 2018.

The eyeWitness to Atrocities App

This is a smartphone camera app that captures metadata to facilitate the authentication of images for use in investigations or trials. It was developed to help human rights investigators take verifiable images. When an investigator uses the app in secure mode to take a photo or video, the app automatically records the time, date and location using the phone's Global Positioning System (GPS) receivers to capture the latitude, longitude, date and time. The app also collects information, where available, about nearby Wi-Fi networks and cell towers, to corroborate the location. It generates a unique alphanumeric code called the hash value for each image. This hash value is calculated using the pixel value, which is based on the pixels' colour, intensity and other factors. If somebody challenges the authenticity of an image taken with the app, the hash value of the original acts like a digital fingerprint that shows that the image has not been edited since it was taken. The location, date and time are stored encrypted with each image to ensure they cannot

be altered. Investigators can send the original image encrypted to the eyeWitness server, where it is securely stored offline. This approach helps to preserve the chain of custody. Those using the app can then save their own copy of the image to use as they wish. The eyeWitness project can certify the image metadata using the original to verify the copy. The app includes a number of security features and is regularly reviewed using feedback.

Collaborations with Investigators

In 2014, independent investigators field-tested the beta version of the app in conflict zones and low infrastructure environments. The eyeWitness app was launched in 2015 as a free download for android devices in the Google Play Store. Since then, the eyeWitness project has collaborated with investigators by providing this free tool for gathering and storing evidence of international crimes for use in reports and litigation. Despite the fact that the project arose out of the crowdsourcing movement, it was determined early on that targeted dissemination was more appropriate. A small number of partnerships with human rights investigators were developed, and the eyeWitness project provided training and ongoing technical support to integrate the app into partners' research methods. Today, the project has more than a dozen partner organisations on four continents. Of the 3000 investigative images submitted to the eyeWitness server, the majority were submitted by partner organisations.

The eyeWitness project leaves the content of investigations to each partner's mandate. While the eyeWitness project can verify the metadata of images taken with the app, it does not investigate or verify the contents of the images or conduct its own advocacy. Instead, the eyeWitness project collaborates with monitoring organisations already conducting human rights investigations, mainly in areas of armed conflict. It aims to support their work primarily by certifying the image metadata and the app technology.

The approach of the eyeWitness project has been to agree to terms on data sharing, security and confidentiality with its partners. Material submitted from the app is securely stored and catalogued in the eyeWitness server, ensuring a protected chain of custody. The project assists its partners' efforts to produce stronger reports in the near term, and eventually to seek justice. Where the eyeWitness project team are not aware

of another group acting on the material to seek accountability and the partner consents, they can seek out credible courts and accountability mechanisms to investigate the images.

Case Study

The following is a case study of a partnership between the eyeWitness project and investigators in Ukraine who agreed to make their use of the eyeWitness app public. International Partnership for Human Rights (IPHR) is an international NGO working with civil society groups primarily in countries of the former Soviet Union, advocating for human rights at the international level. IPHR agreed to use the eyeWitness app as part of their documentation activities undertaken in collaboration with Truth Hounds, a Ukrainian based NGO that aims to fight impunity by recording and monitoring human rights violations and international crimes. Prior to using the eyeWitness to Atrocities app, the Truth Hounds team already collected photos and videos to document alleged violations of international law.

The investigators used a combination of different investigative methods to collect information about the impact of shelling in Eastern Ukraine for reports and submissions to courts. Their research methods include taking witness statements, notes, using Google Earth and Google Maps and verifying information from other field researchers and social media posts. Truth Hounds decided to use the eyeWitness app as an additional tool to verify data collected by their monitors. They did so as part of a "nine-step" model of data collection and verification that aimed to strengthen evidence and reduce opportunities for data manipulation. Their main goal in using the app was to confirm shelling sites in the Donbas region of Ukraine to help reveal whether objects protected under international law such as schools, hospitals, churches and civilian houses may have been targeted and by whom. Truth Hounds took photos and videos of shell casings, craters, property damage and the area surrounding them. They then used metadata captured by the eyeWitness app and other tools to help produce a visual map. The map accompanied the written report "Scorching Winter 2016-2017: Analysis of the shelling of settlements in eastern Ukraine" (IPHR et al. 2017). Truth Hounds did not use data from the app to collect standalone evidence, but instead used it to corroborate and complement information using multiple tools, methods and sources.

Challenges to Consider

Although technology tools have the potential to assist human rights investigators, it is clear that they are not a panacea (Tuckwood 2014; Raymond and Sandvik 2017; Aronson 2016, 2017). The organisations that eyeWitness works with continue to use technology tools as part of a combination of research methods. Our case study supports the view that technology tools are best deployed through collaborative, reciprocal partnerships where assumptions and design are open to consultative re-evaluation (Heinzelman and Meier 2013, pp. 125, 136; The Engine Room 2016b, p. 16).

Technology tools designed for human rights investigations are not suited to all contexts. When deciding on research methods and tools, investigators should consider a number of important issues. What physical and digital security risks will the investigators face? For example, how safe is it to have a smartphone? The availability of equipment, the level of technology skills, access to Wi-Fi, mobile data and cell tower coverage are key factors. The storage and protection of sensitive data is also important with options for paper archiving, electronic databases, encryption, onsite or remote repositories. Witness interviews and statements, survey questions and written reports are used to document human rights violations after events have taken place. Photos, videos, satellite images, apps capturing or protecting metadata, election-monitoring apps, or hidden cameras may offer more flexibility for documenting events as they happen or in the aftermath (The Engine Room 2016a, b; Carter Center ELMO).

Informed consent is a matter of particular concern where technology tools are used as they can create distance between investigators, victims and those acting on the data. Information about human rights violations has serious security, confidentiality and legal privilege implications. However, smartphone technology has made it easy to capture and share images, often without the knowledge of the individuals appearing in them. Human rights investigators should adhere to the "do no harm" principle, which is the ethical obligation not to jeopardise the safety of victims and witnesses (Land et al. 2012; Land 2016, p. 415; UNOHCHR 2011, p. 4). Investigative journalists similarly protect the anonymity of sources in danger as a pillar of the right to freedom of expression (*Goodwin v United Kingdom* ECHR 1996-II). In this context, "do no harm" implies that the person collecting the

data, the victims and bystanders appearing in images should have the opportunity to consent to use of the data where possible and appropriate. The eyeWitness project has sought to navigate these issues by seeking consent on use of the data from the partner organisations and individuals submitting images to the eyeWitness server where they opt to provide contact details.

The use of technology to document violations further complicates informed consent by allowing anonymous data sharing. The human rights investigators who receive information may not know the source of the images to obtain informed consent. This challenge is particularly acute when data collection is crowdsourced. In this situation, not only may the sources be anonymous, they may also lack experience in documenting human rights violations, and therefore may not recognise the risks to those they film or the need to obtain consent. A more targeted approach to data collection like crowdseeding may ameliorate these concerns to an extent as it allows for selection and training of the documenters (Tuckwood 2014, p. 83; Van der Windt and Humphreys 2014, p. 3).

Conclusion

The need to authenticate images posted to social media when ordinary bystanders and perpetrators take incriminating photographs is likely to continue.[4] It is important that technology tools for human rights investigations complement, rather than replace, other documentation and verification methods. Despite the challenges of new technologies, they have the potential to help combat problems posed by "fake news", propaganda and misinformation by helping authenticated images to rise above the noise. Where human rights investigators set out to collect photographic evidence of international crimes for the dual purpose of reporting and litigation, tools like the eyeWitness app can increase their long-term impact for justice and accountability.

[4]See Irving, E. (2017) on the use of social media videos as evidence in an ICC arrest warrant for Mahmoud Mustafa Busayf Al-Werfalli, in the context of the Libya situation.

References

Allan, S. (Ed.). (2017). *Photojournalism and Citizen Journalism: Co-operation, Collaboration and Connectivity*. London: Routledge.

Aronson, J. D. (2016). Mobile Phones, Social Media and Big Data in Human Rights Fact-Finding: Possibilities, Challenges and Limitations. In P. Alston & S. Knuckey (Eds.), *The Transformation of Human Rights Fact-Finding* (pp. 441–462). Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190239480.003.0021.

Aronson, J. D. (2017). Preserving Human Rights Media for Justice, Accountability, and Historical Clarification. *Genocide Studies and Prevention: An International Journal, 11*(1), 82–99. https://doi.org/10.5038/1911-9933.11.1.1441.

Ashouri, A., Bowers, C., & Warden, C. (2014). The 2013 Salzburg Workshop on Cyber Investigations: An Overview of the Use of Digital Evidence in International Criminal Courts. *Digital Evidence and Electronic Signature Law Review, 11,* 115–127. https://doi.org/10.14296/deeslr.v11i0.2130.

Bell, F. (2015, June 16). Verification: Source vs. Content. *First Draft*. Retrieved from https://medium.com/1st-draft/verification-source-vs-content-b67d6eed3ad0.

Bell, F. (2017, April 25). Here's a List of Initiatives That Hope to Fix Trust in Journalism and Tackle "Fake News". *Medium*. Retrieved from https://medium.com/@ferg/heres-a-list-of-initiatives-that-hope-to-fix-trust-in-journalism-and-tackle-fake-news-30689feb402.

Bob, C. (Ed.). (2008). *The International Struggle for New Human Rights*. Philadelphia: University of Pennsylvania Press.

Brandtzaeg, P. B., Lüders, M., Spangenberg, J., Rath-Wiggins, L., & Følstad, A. (2016). Emerging Journalistic Verification Practices Concerning Social Media. *Journalism Practice, 10*(3), 323–342. https://doi.org/10.1080/17512786.2015.1020331.

Carpenter, C. R. (2010). Governing the Global Agenda: Gate-Keeping and Issue Adoption in Transnational Advocacy Networks. In D. Avant, M. Finnemore, & S. Sell (Eds.), *Who Governs the Globe?* (pp. 202–237). Cambridge: Cambridge University Press.

Carter Center. (n.d.). *Election Monitoring App (ELMO)*. Retrieved from http://electionstandards.cartercenter.org/tools/elmo/.

Cohen, S. (2001). *States of Denial: Knowing About Atrocities and Suffering*. Cambridge: Polity Press.

Constable, P. (2007, February 18). Demise of the Foreign Correspondent. *Washington Post*. Retrieved from http://www.washingtonpost.com/wp-dyn/content/article/2007/02/16/AR2007021601713.html.

Fabijanić, D., Spahr, C., & Zlatarsky, V. (Eds.). (2016). *Conflict Reporting in The Smartphone Era: From Budget Constraints to Information Warfare*. Retrieved from Konrad-Adenauer-Stiftung website: http://www.kas.de/wf/doc/kas_47078-544-1-30.pdf?161128161223.

Freeman, L. (2018). Digital Evidence and War Crimes Prosecutions: The Impact of Digital Technologies on International Criminal Investigations and Trials. *Fordham International Law Journal, 41*(2), 283–355.

Golder, Y., & Reich, Z. (2017, May 1). Journalistic Evidence: Cross Verification as a Constituent of Mediated Knowledge. *Journalism, 18*(5), 558–574. https://doi.org/10.1177/1464884915620268.

Goldsmiths University, London. (2011). *Forensic Architecture Project*. Retrieved from http://www.forensic-architecture.org/.

Goodwin v United Kingdom ECHR 1996-II.

Heinzelman, J., & Meier, P. (2013). Crowdsourcing for Human Rights Monitoring: Challenges and Opportunities for Information Collection and Verification. In J. Lannon & E. F. Halpin. (Eds.), *Human Rights and Information Communication Technologies: Trends and Consequences of Use* (pp. 123–138). Hershey, PA: IGI Global.

Hermida, A. (2015). Filtering Fact From Fiction: A Verification Framework for Social Media. In L. Zion & D. Craig (Eds.), *Ethics for Digital Journalism* (pp. 59–73). New York: Routledge.

International Criminal Court. (2004). *E-Court Protocol from ICC OTP Regulations ICC-BD/05-01-09 Regulations of the Court ICC-BD/01-01-04*. Retrieved from https://www.icc-cpi.int/NR/rdonlyres/B920AD62-DF49-4010-8907-E0D8CC61EBA4/277527/Regulations_of_the_Court_170604EN.pdf.

International Criminal Court. (2009). *E-Court Protocol from ICC OTP Regulations ICC BD/05 01 09*. Retrieved from https://www.icc-cpi.int/NR/rdonlyres/…/ICCBD050109ENG.pdf as amended by Prosecutor v. Callixte Mbarushimana, Case No. ICC-01/04-01/10, Decision amending the e-Court Protocol, 4 (April 28, 2011). Retrieved from https://www.icc-cpi.int/pages/record.aspx?uri=1064629.

International Partnership for Human Rights (IPHR), Truth Hounds, & Civic Solidarity. (2017). *Scorching Winter 2016–2017: Analysis of the Shelling on Settlements in Eastern Ukraine*. Retrieved from http://truth-hounds.org/wp-content/uploads/2017/09/last-UA-eng-20.09-web.compressed.pdf.

Irving, E. (2017, August 17). And So It Begins… Social Media Evidence in an ICC Arrest Warrant. *Opinio Juris*. Retrieved from http://opiniojuris.org/2017/08/17/and-so-it-begins-social-media-evidence-in-an-icc-arrest-warrant/.

Koettl, C. (2017). Sensors Everywhere: Using Satellites and Mobile Phones to Reduce Information Uncertainty in Human Rights Crisis Research. *Genocide Studies and Prevention: An International Journal, 11*(1), 36–54. https://doi.org/10.5038/1911-9933.11.1.1440.

Land, M., Meier, P., Belinsky, M., & Jacobi, E. (2012, November). *#ICT4HR Information and Communication Technologies for Human Rights*. World Bank Institute, Nordic Trust Fund, Open Development Technology Alliance & ICT4Gov.

Land, M. K. (2016). Democratizing Human Rights Fact-Finding. In P. Alston & S. Knuckey (Eds.), *The Transformation of Human Rights Fact-Finding* (pp. 399–424). Oxford: Oxford University Press.

Lowe, R. (2015, June 11). *eyeWitness: Witnessing Atrocity*. International Bar Association. https://www.ibanet.org/Article/NewDetail.aspx?ArticleUid=11e76b66-d949-4738-9347-e67fbfbb9441.

Mast, J., & Hanegreefs, S. (2015). When News Media Turn to Citizen Generated Images of War: Transparency and Graphicness in Visual Coverage of the Syria Conflict. *Digital Journalism, 3*(4), 594–614. https://doi.org/10.1080/21670811.2015.1034527.

McPherson, E. (2015a). *ICTs and Human Rights Practice: A Report Prepared for the UN Special Rapporteur on Extrajudicial, Summary, or Arbitrary Executions. Centre of Governance and Human Rights*. Cambridge: University of Cambridge.

McPherson, E. (2015b). Digital Human Rights Reporting by Civilian Witnesses: Surmounting the Verification Barrier. In R. A. Lind (Ed.), *Producing Theory in a Digital World 2.0: The Intersection of Audiences and Production in Contemporary Theory (2)* (pp. 193–209). New York: Peter Lang.

McPherson, E. (2016). Source Credibility as Information Subsidy: Strategies for Successful NGO Journalism at Mexican Human Rights NGOs. *Journal of Human Rights, 15*(3), 330–346. https://doi.org/10.1080/14754835.2016.1176522.

McPherson, E., & Probert, T. (2017). Special Procedures in the Digital Age. In A. Nolan, R. Freedman, & T. Murphy (Eds.), *The United Nations Special Procedures System* (pp. 261–270). Leiden, Boston: Brill Nijhoff.

Pantti, M., & Sirén, S. (2015). The Fragility of Photo-Truth: Verification of Amateur Images in Finnish Newsrooms. *Digital Journalism, 4*(3), 495–512. https://doi.org/10.1080/21670811.2015.1034518.

Platt, E. (2014, October 9). Citizen Journalists Playing a Crucial Role in Syrian War. *Time*. Retrieved from http://time.com/3481790/syria-journalism-kobani/.

Powers, M. (2015). NGOs as Journalistic Entities: The Possibilities, Promises and Limits of Boundary Crossing. In M. Carlson & S. C. Lewis (Eds.), *Boundaries of Journalism: Professionalism, Practices and Participation* (pp. 186–199). New York: Routledge.

Raymond, N. A., & Sandvik, K. B. (2017). Beyond the Protective Effect: Towards a Theory of Harm for Information Communication Technologies in Mass Atrocity Response. *Genocide Studies and Prevention: An International Journal, 11*(1), 9–24. https://doi.org/10.5038/1911-9933.11.1.

Reporters Without Borders (2013, November). *Journalism in Syria, Impossible Job?* Retrieved from https://rsf.org/en/reports/journalism-syria-impossible-job.

Safran, S. (2005, December 15). *How Participatory Journalism Works* (Nieman Reports). Retrieved from http://niemanreports.org/articles/how-participatory-journalism-works/.

Sambrook, R. (2010, December). *Are Foreign Correspondents Redundant? The Changing Face of International News*. Oxford: Reuters Institute for the Study of Journalism & Oxford University Press. Retrieved from http://reutersinstitute.politics.ox.ac.uk/publication/are-foreign-correspondents-redundant.

Shapiro, I., Brin, C., Bédard-Brûlé, I., & Mychajlowycz, K. (2013, December 1). Verification as a Strategic Ritual. *Journalism Practice, 7*(6), 657–673. https://doi.org/10.1080/17512786.2013.765638.

Silverman, C. (2012). *A New Age for Truth* (Nieman Reports). Retrieved from http://niemanreports.org/articles/a-new-age-for-truth/.

The Engine Room. (2016a). *Datnav: How to Navigate Digital Data for Human Rights Research*. Retrieved from https://www.theengineroom.org/wp-content/uploads/2016/09/datnav.pdf.

The Engine Room. (2016b). *Technology Tools in Human Rights*. Retrieved from https://www.theengineroom.org/wp-content/uploads/2017/01/technology-tools-in-human-rights_high-quality.pdf.

Tuckwood, C. (2014). The State of the Field: Technology for Atrocity Response. *Genocide Studies and Prevention: An International Journal, 8*(3), 81–86. https://doi.org/10.5038/1911-9933.8.3.7.

Umberg, T., & Warden, C. (2014). The 2013 Salzburg Workshop on Cyber Investigations: Digital Evidence and Investigatory Protocols. *Digital Evidence and Electronic Signature Law Review, 11,* 128–136. https://doi.org/10.14296/deeslr.v11i0.2131.

United Nations Office of the High Commission for Human Rights (UNOHCHR). (2011). *Basic Principles of Human Rights Monitoring Manual on Human Right Monitoring*. No. 7/Rev. 1, HR/P/PT/7/Rev. 1. Retrieved from http://www.ohchr.org/Documents/Publications/Chapter02-MHRM.pdf.

Van der Windt, P., & Humphreys, M. W. (2014). Crowdseeding in Eastern Congo. *Journal of Conflict Resolution, 60*(4), 748–781. https://doi.org/10.1177/0022002714553104.

Wardle, C., Dubberley, S., & Brown, P. (2014). *Amateur Footage: A Global Study of User-Generated Content in TV and Online News Output*. Retrieved from the Tow Center for Digital Journalism website: http://towcenter.org/wp-content/uploads/2014/05/Tow-Center-Amateur-Footage-A-Global-Study-of-User-Generated-Content-in-TV-and-Online-News-Output.pdf.

Witness. (n.d.). *Video as Evidence Field Guide*. Retrieved from https://vae.witness.org/video-as-evidence-field-guide/.

Check for updates

CHAPTER 15

A Matter of Perspective: Truth, Evidence and the Role of Photography as an Investigative Tool

CJ Clarke

"When I use a word," Humpty Dumpty said in rather a scornful tone, "it means just what I choose it to mean — neither more nor less."

Lewis Caroll, Through the Looking Glass

It is said, the camera never lies. The camera always lies. Truth, evidence, and the whole photographic tradition of bearing witness is really only a matter of perspective. Whose truth, and whose pain, to paraphrase Susan Sontag, do we choose to regard?

The role of photographer, image maker, or artist—all essentially interchangeable terms—is opaque: What they choose to frame, is dependent on the process of exclusion. Stephen Mayes notes that "there is a vast gulf between facts and truth, and history is strewn with factual images

CJ Clarke (✉)
JAPC, London, UK

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_15

that have been used to tell lies. So, in short, photography's relationship with truth is little more than 'belief' imposed by the viewer."[1]

Further, we must also consider the media in which visual images appear or are deemed fit to appear. Where and how images are used is an important controlling factor in determining 'whose' truth is told, and crucially, 'how' it is told. Traditionally, the individual practitioner does not control the means of production, for instance, the headline associated with a particular image and perhaps not even the caption. All in all, the narrative is shaped by editors far removed from the circumstances of the images making. Without specific expertise, images are often viewed through the prism of an editor's own cultural understanding, which invariably is an extension of their publications politics.

In this regard, we begin to understand the layers of framing taking place in the photographic rendering of reality. First, there is the photographer's own methodology, personal ethics and politics all leading to the 'actual' act of framing—that is the taking of the photograph; and, secondly, an act of framing occurs in how publications use and contextualise a particular image or set of images. It is not inconceivable that a photographer's own motivation or understanding of the situation is at odds with the wider editorial position of a particular publication or the contextual information provided by a particular publication to sit alongside their images.

So, how does all of this relate to investigative journalism? Unlike other forms of investigative journalism—balance-sheet investigations, for instance—the subjectivity of photographic representation makes this a tricky question to answer. Indeed, within the confines of this article it is, perhaps, an impossible question to answer. At its root is the inherent contradiction that photography *is* the representation of reality and is simultaneously *not* the representation of reality. Such a duality keeps us on the forked path.

For the purpose of this article I will, initially at least, focus on the genre of photojournalism, as traditionally this is the area of photography that is most closely aligned with investigative journalism. Conceding that I will be unable to provide any concrete 'answers' to the use of photography within an investigative context, I will, instead, ruminate,

[1] Stephen Mayes, interview with the author.

playfully and provocatively, on where we have been and how we might move forward in the future.

For photojournalist Nic Dunlop, "of course, photography is the truth,"[2] but he acknowledges its limitations, as it contains an inherent "duality" that is not "widely understood or generally acknowledged by practitioners, particularly in the world of photojournalism. There is still the belief that a photograph 'documents the truth' which, on one level, is true in and of itself. But the major flaw of this his view is that it implies 'objectivity', as though there is some nirvana of neutrality that can be achieved."[3]

Dunlop continues to note "the idea that a photograph can contain many truths simultaneously is still hard for many to accept, especially if they are contradictory truths." He cites the example of the photographs of those incarcerated in the infamous Khmer Rouge prison, S-21: "Many of those incarcerated were Khmer Rouge themselves. Some were killers who fell out of favour with the regime and found their roles reversed and we are faced with images of victims who may have been perpetrators. It complicates our view."[4]

Dunlop contends that it is not "photography as a medium that is the problem, but the way we choose to engage with a complex medium. Photography can assist in helping us to understand on an emotional level, but it can't explain. Now, with smartphones even more stop to think about the images and much fewer understand their power and why some images arrest and others fail to engage."[5] Similarly, filmmaker Peter Greenaway has often commented on the inadequacies of visual imagery in describing the representations of specific instances. Visual imagery excels in the communication of atmosphere and mood—in all that is non-verbal. Allied to the subjectivity inherent in the photographic gaze, we can further glimpse the uneasy relationship between the photographic object and notions of "evidence" or, indeed "truth."

All of this is another way of saying that the photographic representation of the world is highly subjective. Whilst the camera can record a moment and provide a 'record' of a particular event, situation

[2] Nic Dunlop, interview with the author.

[3] Nic Dunlop, interview with the author.

[4] Nic Dunlop, interview with the author.

[5] Nic Dunlop, interview with the author.

or person(s), the validity of this record as unadulterated reality is questionable.

And yet, photography`s power to capture our imagination is unparalleled. It is a galvanising force, pulling together disparate arguments into one readily understandable visual reference, shaping our view of events, situations and people like no other. However, the ability of still images to do so is dependent on its distribution through the mass media and it would be possible to argue, perhaps cynically, that in this context, photography simply provides a visual accompaniment to the predominant narratives that are 'fit to print' according to our mainstream news outlets. As Nic Dunlop notes, "We want our stories told back to us, to have our prejudices confirmed. And we tell the same stories over and over again to shore up our own values."[6]

But who believes the news? Given the plurality of twenty-first century media outlets and, especially in the 'post-truth' era and a steady flow of disinformation, there is perhaps a wider, popular critique of what images can and do (or do not) represent. Whilst scepticism is healthy, the trend towards viewing images (or indeed any news) is worrying. One group where belief in the power of images remains strong, however, is within the photojournalistic industry itself. This belief, in many cases, is rigid and built on foundations of a mythical set of values, located somewhere in the industries mid-twentieth century heyday. (For 'heyday', read the time when photojournalism was well funded and had access to a large audience through photo-driven magazines such as *Life* and *Picture Post* and had not lost its audience to television news.)

From this time, the dominant form to emerge was the photographic essay. This collection of twelve to twenty images published initially in Life and similar publications and latterly in the glossy magazines found in weekend newspapers probably represents the chief way in which most people interacted with reportage and documentary visual content until relatively recently. The leading proponent of the photo essay to emerge in the 1950s was the photographer W. Eugene Smith. He is an iconic figure within photojournalism, lionised for his dedication to both his craft and the stories he chose to tell, most notably his long-term commitment to the story of the Minamata chemical disaster in Japan.

[6] Nic Dunlop, interview with the author.

Smith's much-quoted line—"With considerable soul searching, that to the utmost of my ability, I have let truth be the prejudice" (Maddow and Smith 1985)—has become something of a mantra for photojournalism, and his position is untouchable. Upon closer inspection, however, Smith's legacy rather more points towards the limitations of photojournalism and any rigid adherence to a singular 'truth' or proscribed ethical standpoint. He is the proto 'photographer-as-activist' and his work does rather more to advance the greater need for the subjective interpretation than his champions would care to admit. In this sense Smith occupies a more radical position that questions the role of mass media in the formation of dominant narratives and the relationship between said media and individual practitioner contained therein. In such a context, the elevation of his personal mantra to industry credo would rather seem to miss the point.

The desire within photojournalism, to investigate by 'bearing witness', with 'truth' on your side, would seem rather a naive standpoint; it emulates Smith's enthusiasm without acknowledging the ambiguities highlighted through his work. As Paul Lowe (2014) identifies, "this role of witness is more complex than at first sight" (p. 227) as it contains certain limitations. He quotes the photographer Susan Meiselas:

> [T]he other side of 'witness' is that we do intervene, and we intervene by the fact of our presence in a particular place. We change how people see themselves sometimes and how others may come to see them. I'm also concerned about how we see ourselves in the process of our role as witnesses. (as cited in Lowe 2014, p. 227)

Meiselas identifies photojournalism's self-image problem, moulding itself as a vanguard 'truth-seeker,' without acknowledging the problems of such a position. Moreover, Meiselas' comments acknowledge that the presence of the photographer influences the scene that they are depicting and, further, that the 'evidence' as gathered by a photojournalist bearing witness is highly partisan. It is of no surprise that Meiselas' principle concern is human rights violations and much of her work—photographic or otherwise—embraces the partisan nature of this approach, very much in the spirit of Smith, with the photographer occupying the role of activist or, at very least, an active participant or visible presence within the story they are trying to tell.

In further reference to the work of Smith we might choose to look at his essay on Franco's Spain—an extremely powerful set of pictures; they are, however, something more than an impartial document of a moment in time. In one memorable image, three Spanish soldiers are photographed looking into the sun. Their faces chiselled in harsh light and shadow, they are ominous, and Smith's own opinion of the dictatorship is revealed. In another famous example, an image of a group mourning a dead family member, Smith doctored the photograph because a woman on the periphery was smiling—not in keeping with the overall tone of the photograph. Smith is partisan and involved, shaping the images based on his institution and opinion. Such involvement—and active doctoring of pictures—would fall foul of many modern 'ethical' standards; nonetheless, few would argue that this work is not an 'investigation' into a particular society, at a particular time. But his findings, his conclusions are nuanced and non-verbal; they might not stand up to scrutiny as 'evidence' but I wonder if they tell us more about this society than a set of images that attempts to be more impartial, distanced and objective?

Over the course of the twentieth century, the role of photography has become intertwined with our ability to accept an event as significant. The act of photographing, however subjective, is evidence not only of an act actually occurring but is a signifier of importance: This event is important because it has been photographed. As Judith Butler notes, photography, with particular reference to conflict and pictures of atrocities, is "built into the notion of atrocity, and photographic evidence establishes the truth of the claim of atrocity in the sense that photographic evidence has become all but obligatory to demonstrate the fact of atrocity" (as cited in Lowe 2014, p. 227). Lowe (2014) further notes how this "is the continual dilemma of the witness of atrocity, the more desperate and debased the story, the more significant it becomes and the more necessary and likely it is to be reported" (p. 228).

This brings us back to the dualism of photography: A photograph is evidence of the real, but that reality is only 'real' because it is photographed. And, as I enquired at the top of this essay, how does this fit within media constructs of what is fit to print or, indeed, how is this photographic evidence handled, understood and contextualised by the largely western institutions that publish this material? What are their roles and responsibilities in 'decoding' photographic evidence and presenting it to an audience?

Stephen Mayes notes how "impartiality and objectivity of news coverage were fantasies promoted by media proprietors as a means of bringing credibility to their advertising platforms (AKA newspapers and magazines)".[7] He continues to note that in the twenty-first century "photographers increasingly identify themselves as commentators, in which capacity they bring judgment as well as supposedly fact-based observation." Mayes' interview with the Indian photographer Poulomi Basu is instructive in this regard. Basu takes a "multidisciplinary" approach. She "inhabit[s] the roles of human rights documentarian, journalist, artist and activist" noting that it is "often others who project limitations or boundaries to an individual practitioner or, indeed, have an extreme desire to impose constrictive labels" (Mayes 2017).

In a startling example of orientalism, Basu's work on hidden and normalised violence against women in sub-continental Asia in her project "*Blood Speaks: A Ritual of Exile*" was censured after publication by *The New York Times* because it presented a more complex and nuanced view of the 'orient'. Basu notes how "the project tackles aspects of the Brahmanical gaze, which is manifest throughout untouchability, gender and caste. Such concepts might be less readily understandable to those outside of the subcontinent. Even in the subcontinent, these issues are considered taboo" (Mayes 2017). Highlighting the power imbalance that exists between individual practitioners and media institutions, Basu notes how, in her case, "the institutions of Western media have culturally appropriated the issues to put themselves in a dominant position so that they may cast judgement from a position of perceived cultural superiority" (Mayes 2017). In this case, Basu's perception is rejected by "institutions that are unwittingly reinforcing the structures of patriarchy and prejudice" (Mayes 2017) because of their inability to understand and engage with a subject from beyond a western viewpoint. The intentions and expertise of the practitioner are sidelined; the photograph as evidence is censured because the western institution lacks the expertise and the cultural humility to assess and understand the evidence as it is presented.

It is worth noting that, in Basu's case, as with many of a new generation of photographic practitioners, the term photojournalist is assiduously avoided. Photojournalist has become a loaded term, associated with a rigidity and sense of cultural superiority that is misplaced in

[7] Stephen Mayes, interview with the author.

the multi-variant and multilayered media (and social) landscape of the twenty-first century. As Fred Ritchin (2010) notes in his book *After Photography*, in the digital age the "photograph may be seen as representing a point of view that is more synthetic and impressionist" (p. 57). He continues to note how in "the digital arena one cannot with any certainty look at a photograph and say, 'So that is how it was'" (p. 58). A photograph becomes a "memory magnet" that "others can link to … amplifying and contravening what its initial author claimed it represented" (p. 59). Thus, a photograph becomes a starting point in a wider dialogue rather than an end, a static piece of 'evidence' as delineated and asserted by 'photojournalists' working within what has become, particularly in the twenty-first century, an increasingly codified and reductive framework.

For Ritchin, we certainly should not "abandon the documentary photograph as a credible recording device" because that would "handicap a democracy's capacity to function due to a dearth of credible evidence," but certainly the "growing inability of many governments and citizens to assimilate and respond to nonlocal events, from global climate change to the mass killings in Darfur, suggests that the kinds and amounts of imagery available are already contributing to a cynical breakdown in governance" (Ritchin 2010, p. 62).

First written in 2009, Ritchin's words, in the context of 'post-truth', are remarkably prescient. Despite the abundance of images to be found in this new 'digital' world, the public are turning away—the value and effectiveness of such work is in question. We may postulate that the overuse of familiar tropes, narratives and an aesthetic associated with traditional photojournalism is, at least in part, contributing to such a situation. If this is the case then, where does this leave us? Where do we go from here?

For Lowe (2014), one answer is to fully recognise the photograph as an "independent artefact in and of itself as well as serving as the visual testimony of the photographer." The photograph is "a visual record of the material world, yet it is also a thing in itself, with its own material qualities" (p. 229). Lowe continues to note how "like shards of a tarnished old mirror that has been shattered, photographs reflect the past into the present in a fragmentary, partial way" (p. 230) and that the move towards photographing landscape and objects—what we might call the secondary act of witnessing—positions modern photojournalism into a more relevant and powerful space. The photographer no longer

needs to be photographing an event as it happened; rather allied to survivor testimony it "becomes an affirmation of the rights of the victim to be known and heard," and the "social act of witnessing" or re-witnessing to an audience "serves to re-suture the victim into the social fabric" (p. 235). Mayes notes how "photography has long been thought to stand for truth simply because of the indexical quality which in the analogue medium imprints fact-based information onto film; but even at its most optimistic the medium is never truly factual when one considers the effects of blur, lens distortion and other physical phenomena."[8] Photography thus, assumes "a symbolic role" making full use of its qualities "to encompass both the referential and imaginative elements" (Lowe 2014, p. 235).

By way of example we may look at the work of photographer Gilles Peress, whose work has progressed from a more classical style (albeit filtered through postmodernism and influenced by the fragmenting visual techniques of Lee Friedlander) to something more akin to a crime scene photographer with "close up images of the still lives of the aftermath of atrocities, a visual trope that establishes this 'legalistic' approach and validates the more conventional photojournalist images in their projects, giving them a deeper meaning than a conventional presentation might obtain" (Lowe 2014, p. 237). Lowe (2014) argues that such a visual strategy "builds up a more complete sense of a depth of coverage and evidence" and the,

> similarities between the role that the photograph plays in journalistic discourse and in the courtroom is [*sic*] striking. In both arenas the photograph serves to help the audience grasp complex issues by simplifying them and by giving a sense of place and orientation that textual descriptions alone cannot deliver. (p. 237)

Embracing a more forensic approach is not, however, the only strategy for forging a new language beyond the historic conventions of photojournalism. As Basu earlier noted, her multidisciplinary practice not only means that she inhabits many roles—artist, photographer, activist—but also that she utilises a variety of techniques, from conventional photography and film through to virtual reality. Similar to Peress' use of forensic photography, Basu's methodology expands the lexicon of

[8] Stephen Mayes, interview with the author.

photography, allowing her work to contain a multiplicity of voices moving her work closer to "an oral tradition where divergent views of the community are taken into account" (Ritchin 2010, p. 58). In embracing such multiplicity, Basu's visual strategies move us, the audience, towards greater immersion and understanding of a situation in all its nuance and complexity without the need for text-based linear explanation. Her work is 'evidence' but rejects the binary—and power imbalance—that this has traditionally meant within the strict limits of photojournalism.

The commonality between these two practitioners suggests that the traditional lexicon of photography is inadequate for the purpose in which it has been traditionally cast. Mayes notes that the "common anxiety is that we need to find ways to shore up the investigative functions of the press even as the rest of the institution moves into new territories. But maybe the concern is misplaced, and we should instead be democratizing investigation."[9]

Mayes pursues this radical idea further to suggest that his "ideal reporting team would comprise a psychologist to structure communications that get through people's defences, a games engineer to create incentives for following stories, and an advertising creative director because they're the people who use imagery to really change people's behaviour".[10] He continues, "In other words I'm not sure that we should even want to sustain any credibility for twentieth century protocols … because these standards had less to do with truth-telling and more to do with white, male capitalists trying to establish credibility for the advertising platforms."

We may even turn our attention away from photography altogether or, at least, photography as it is commonly understood. Three-dimensional scanning, for example, is a rapidly developing technology that has the power to record far more information than is possible with any still or movie camera. In essence, it creates an enormous data set of a particular environment which can be recreated as a three-dimensional model or two-dimensional rendering, from whatever perspective the 'author' chooses. Beyond recording the shape of a certain environment, it can also record data on its appearance. The UK company ScanLab

[9] Stephen Mayes, interview with the author.

[10] Stephen Mayes, interview with the author.

claims that it is "a form of machine vision that we argue is the future of photography" (ScanLab, n.d.-a).

Scanning is already being utilised for forensic purposes. A collaboration with the London Fire Brigade has seen the scene of a fire scanned and recreated so that the source of fire may be identified. Or, as in their collaboration with *The Guardian*, "Crime Scene" (ScanLab, n.d.-b), the technology is used to create a docu-fiction "interactive experience" that "brings to life the fascinating world of forensic science, allowing you to experience and understand the crucial role it plays during an investigation" (ScanLab, n.d.-b). Nevertheless, the technology's ability to render reality, in full 360 degrees, to a degree that allows it to be used for forensic purposes perhaps questions the role that photography holds as one of the chief arbiters of visual 'evidence.'

A further glimpse into the possibilities afforded by this technology is demonstrated by an early experiment by a team led by Marc Levoy at Stanford University. The team scanned a number of Michelangelo sculptures with such accuracy that it was possible to see Michelangelo's chisel marks (Levoy, n.d.).

This is 'pure documentary' where a data set is collected free from the hindrance of any bias or perspective; free from the limitations of a two-dimensional camera; free from the limitations of what an individual is able to capture from one point within a location. Here, everything is collected, and it is only later that this moment is recreated so that it can be interrogated from multiple perspectives.

Similarly, Forensic Architecture (FA) "is a research agency, based at Goldsmiths, University of London, that undertakes advanced architectural and media research on behalf of international prosecutors, human rights organisations and political and environmental justice groups" (Forensic Architecture, n.d.). Within this emergent field, 3D scanning is a tool to be used alongside the collection of other data, notably film and photos captured by those actually being affected by the event under investigation. Unlike previous decades, the development of technology, notably in the form of smartphones, has cast the 'victims' in the role of their own documentarians. We no longer have to regard the pain of others, we can regard our own pain.

As Mayes notes, this is the "reading space to reconstruct behaviors, or to validate people's description of events."[11] For FA "analysing IHL [International Humanitarian Law] and HR [Human Rights] violations must involve modelling dynamic events as they unfold in space and time and creating navigable 3D models of environments undergoing conflict, as well as the creation of filmic animations, and interactive cartographies on the urban or architectural scale" (Forensic Architecture, n.d.).

A look at the history of arts tells us that the development of a new technology, more suited to the portrayal of reality, liberates its antecedents to explore more abstract and nuanced territory. This was the case for painting with the invention of photography, for instance. Such developments are, perhaps, the liberation of photography from a role in which it was ill cast. As Nic Dunlop states, "We want the world simplified and reduced to simple formulas of 'good' and 'evil' … we are fixated on this interpretation in order to bring some kind of order to the messy and grim realities of our world."[12] Whilst understandable, this response can be "dangerous as it shores up prejudices based on feeling alone rather than contribute to a greater understanding."[13]

Photography is the truth, in a sense, but it is also so much more. Greater awareness, by both practitioners and audience, of its limitations as 'evidence' is also an awareness of its strengths. At less than 200 years old, photography is still a new art form, especially when compared with the vast lineage of painting and literature. The photographic language will continue to mutate. It will do so faster than those who make facile and arrogant attempts to codify its language and proscribe its ethical standpoint. It will keep evolving and shifting; we must keep developing and interrogating so that we can, to paraphrase T.S. Eliot, arrive where we started and know that place for the first time.

References

Forensic Architecture. (n.d.). *Project*. Retrieved from http://www.forensic-architecture.org/project/.

Levoy, M. (n.d.). *The Digital Michelangelo Project*. Retrieved from http://graphics.stanford.edu/projects/mich/.

[11] Stephen Mayes, interview with the author.

[12] Nic Dunlop, interview with the author.

[13] Nic Dunlop, interview with the author.

Lowe, P. (2014). The Forensic Turn: Bearing Witness and Thingness of the Photograph. In L. Kennedy & P. Caitlin (Eds.), *The Violence of the Image: Photography and International Conflict* (pp. 226–254). London: I.B. Tauris.

Maddow, B., & Smith, W. E. (1985). *Let Truth Be the Prejudice; W. Eugene Smith, His Life and Photographs*. New York, NY: Aperture.

Mayes, S. (2017, July 19). Poulomi Basu on Her Project "Ritual Of Exile"~In Discussion with Stephen Mayes. *Tim Hetherington Trust*. http://www.timhetheringtontrust.org/news-and-calendar/2017/07/poulomi-basu-her-project-ritual-exile-discussion-stephen-mayes.

Ritchin, F. (2010). *After Photography*. New York, NY: W. W. Norton & Company.

ScanLab Projects. (n.d.-a). [Info page]. Retrieved from https://scanlabprojects.co.uk/about/.

ScanLab Projects. (n.d.-b). *Crime Scene*. Retrieved from https://scanlabprojects.co.uk/work/crime-scene/.

Check for updates

CHAPTER 16

The Context Verification of Digital Photographs

Alexander Godulla

Introduction

It is nearly impossible to determine the number of photographs taken every day. As a direct result of the digitalisation of photography and the rise of the smartphone, taking pictures has become an almost everyday routine for billions of people. In a country like Germany alone, eleven percent of all the photographs created this way enter social networking sites within 60s (Godulla 2014, p. 402). Even without collecting comparable studies from other parts of the world, it is obvious that these shared activities are increasing in strength. For journalists, this phenomenon is a blessing and a curse at the same time. A blessing, because it has never been that easy to gather visual material of public interest. A curse, because the material provided can often be the subject of manipulation. As the essay of Tanasi in this volume shows, tampering can be identified by using certain kinds of software. This technique of verification often goes along with an ethic discussion. One of the most popular examples is concerning the World Press Photo of 2013, which created its impact on the audience by using an almost unreal use of shadow and light.

A. Godulla (✉)
Institute of Communication and Media Studies, University of Leipzig, Leipzig, Germany

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_16

It was taken on November 20, 2012, showing dead children who are carried through the streets of Gaza City after an Israeli attack carried out as retaliation to Palestinian rocket fire. An extensive software analyses showed traces of intense postproduction (especially dodging, burning), but also proved that no pixels had been moved during this process (Anthony 2013). The picture was taken by Paul Hansen, a photographer of Danish origin. He remembers that he had to defend his image against numerous attacks (Gailītis 2013):

> Then the usual criticism came – it is too dark, too toned, too post-processed and then again I defended and said what I did and did not. The attacks have come from many different corners […] Finally enough was enough and the World Press Photo sent the raw file to 3 different experts – two in the US and one in the Netherlands. […] They went pixel by pixel and I was cleared for any manipulation which was really good because after that it stopped.

Hansen also explains the persistent discussion about his photograph with the fact that its content is politically inconvenient: "Of course, I realize that it is an uncomfortable picture for many people, it has a very political context. Many people don't like it, especially political players, because it is a manifestation of a political failure" (Gailītis 2013). But the discussion, where the enhancement of photography in general ends and where fake photography begins could not be decided then and is still open to debate. In contradiction to that, this essay focuses on aspects of photography that can be described as "true" or "false", leading to questions like:

- Who or what is the subject of the picture?
- Where has the photograph been taken?
- When did the events pictured occur?
- What is the context of the photograph, the bigger picture?

Although an important focus of the authors work lies on new formats in journalism (Wolf and Godulla 2016), this contribution will not take a theoretical perspective. Instead, it aims to give its reader a practical description of the context verification of digital photographs. This technique is mainly based on journalistic research and fuelled by brainpower. The strategy provided is the direct result of many practical courses conducted by the author in journalism studies, where students were

confronted with similar questions. Another important inspiration can be found in a Tipsheet based on a workshop by the German journalism society Netzwerk Recherche (Bremer 2017). The combination of these elements create the foundation for Fig. 16.1: Aspects like the source and the seriality of a photograph, the quality of data, methods of external validation, conspicuous features, the analyses of similar pictures and the plausibility itself are combined in order to create a checklist consisting of seven steps.

These steps will be described in detail in order to offer the reader an easy and effective tool for the process of context verification.

Step 1: Source

In order to assess the credibility of a photograph, the quality of the source should be taken into consideration first. Journalists should always prefer sources of particularly high credibility. In addition to relevant news agencies, these are also known and professional photographers, who are usually working for the editorial staff on a regular basis. If the origin of a photograph cannot be clarified, it should not be spread. This is particularly the case for images from social networking sites. Frequently, it is not the images themselves, which are problematic. It is rather the caption, which is often deliberately (to manipulate people) or mistakenly (in ignorance) falsified, changing the meaning of a photograph by misinterpreting its context. Social networking sites should therefore be regarded as a toxic environment for any form of image research. Trolling, flaming and hate speech are the order of the day, especially when dealing with controversial subjects, which makes the assessment of the authenticity of the used imagery extremely difficult.

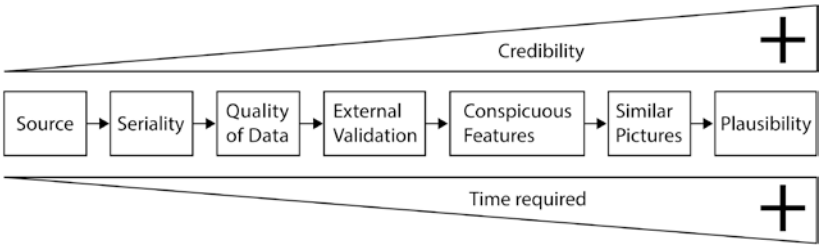


Fig. 16.1 Aspects of context verification in press photography

Step 2: Seriality

In the analogous age, many popular cameras typically used 36 slides of negative film per roll. Accordingly, it was costly and tedious in the past to make picture serials. In any case, photographers are always looking for the decisive moment, that is, the situation in which form and content are in the best possible relationship. The photos taken before and after this one shot are generally not published and are thus irrelevant to the public. This has not changed in the digital age. However, it was common then and is common now to create larger numbers of photographs when being confronted with highly relevant motives. For example, a modern professional photojournalist would take more than 1000 photographs during an important soccer game (Godulla 2009, p. 50). Even with portrait photos, many photographers today use the serial mode of a camera in order to capture the best possible facial expression of a person. When assessing the authenticity of a possibly faked photograph, it is therefore a good strategy to ask the source for further material from the same series. Frequently, additional aspects of the situation become visible, which allow a better assessment of the image. If the source cannot or will not provide further material, the credibility of the justification is a further hint for the authenticity of a photograph.

Step 3: Quality of Data

In the age of digital photography the use of film has been massively pushed back. Instead, photos are stored in certain data formats, each with certain properties. So-called JPEGs provide a lossy compression for digital images. This format is used by nearly every amateur camera, including smartphones. Even professionals often rely on JPEGs when having to deliver photos to professional customers like editorial offices. However, JPEGs provide only a little potential for shadow recovery or highlight recovery. Legitimate and traditional techniques of image optimization are thus made more difficult using them. Professional photographers therefore often rely on the so-called RAW format, which creates a kind of digital negative. This file requires much more disk space, but also allows much better post-processing. Therefore, the majority of professional photographers tend to use both formats in their workflows. Digital manipulations are much more difficult to apply to RAW-files than to JPEGs. For this reason, editorial teams should request

access to any RAW-file related to a photograph of questionable authenticity. In addition, the accompanying metadata should also be included in the assessment. Although Metadata can easily be falsified, it is a possible source of error for any counterfeiter. Metadata can, for example, list the GPS coordinates of the location of the photograph, the focal length of the lens used, or the shutter speed. These values should be compared with the caption of the picture already provided. A programme being able to read metadata (such as Adobe Lightroom) should be used for this purpose. Metadata readers available only online should be avoided in the meantime, since particularly sensitive image material should not be uploaded to the World Wide Web without further need.

Step 4: External Validation

Whether this fourth step is necessary at all depends on the course of the previous analysis steps. If the source is credible (step 1), if the photo can be interpreted as part of a larger series (step 2) and if the data is of high quality (step 3), a fair amount of journalistic care has already been provided. If, on the other hand, the photo does not meet all or none of the criteria, it must be regarded to be very questionable. In this case, an editorial office must ask whether such a photo should be distributed at all. If there is a compelling reason for this, the search radius should now be extended to the World Wide Web. The use of monitoring pages such as snopes.com, which deal with lies spread in cyberspace, may be helpful. These sites allow searching for keywords or even URLs. However, this requires that the relevant topic has already been exposed as a hoax. Similarly, search engines like google.com can be used to find comparable clues through a keyword search. If this is unsuccessful, the photo is far from genuine: it can also only be an indication that you do not use the adequate keywords.

Step 5: Conspicuous Features

If no suspicious fakes can be found elsewhere, it is worth taking a closer look at the photograph on the actual image plane. When following this strategy, it is important to free yourself from your own prejudices as a viewer: how a refugee, a soldier or a passer-by can look like will be very different, depending on the context. In the age of globalisation it is, for example, completely normal that people wearing printed T-shirts appear

in tribal cultures. Nevertheless, it is worth taking a look at the attire of people who are depicted: does it, for example, correspond to the cultural context, religious traditions or climatic conditions? If the photo is sufficiently important, experts with relevant knowledge can be consulted here. Similarly, it is worth searching for everything that contains writing. Road signs, number plates or advertisements are a good example. In this case too, people with regional knowledge can be consulted in order to identify the culture and the country. Strictly speaking, any element on the picture can serve as an indication: Power sockets, for example, indicate which standard is used in a country. Floor coverings can massively differ by country, making them correspondingly meaningful. The same applies to shadows or the position of the sun in the sky, which can be helpful when identifying the time of recording.

Step 6: Similar Pictures

This step is always effective when a photograph has been published already in this form or a very similar form to the World Wide Web. In this case, you can use the image search images.google.com to insert the URL on which the image is to be found. Alternatively, you can also upload the image directly to find it elsewhere. The context published here can be very helpful in assessing the authenticity of the image and to get further clues—also for the further external validation (step 4) or the analysis of conspicuous features (step 5).

Step 7: Plausibility

A final step in the assessment of a photo can refer to the general plausibility. Visible phenomena (such as the weather) are compared with databases. The wolframalpha.com search engine makes it possible to search historical weather data in a very comfortable way. So if there is sunshine on a photo and the database indicates heavy rain, you might be confronted with a fake.

Conclusion

With the help of the seven steps described, it is comparatively easy to obtain a robust impression of the authenticity of a photograph. In journalism, however, it is important to consider not only the search for truth

but also factors such as time and costs. In particular the search for conspicuous features (step 5) can devour a lot of time, if one would like to proceed with the necessary thoroughness. If pictures are not even plausible based on the review of their sources (step 1), they may not be worth this particular effort. Visual communication is always accompanied by a great risk: what people see with their own eyes, they automatically perceive as true. At this moment, photographs become the object of interpretation and possible further spreading. Perhaps the most sensible response to this problem may be the decision not to spread questionable picture material at all.

References

Anthony, S. (2013, May 13). *Was the 2013 World Press Photo of the Year Faked with Photoshop, or Merely Manipulated?* [Web log post]. Retrieved from http://www.extremetech.com/extreme/155617-how-the-2013-world-press-photo-of-the-yearwas-faked-with-photoshop.

Bremer, A. (2017). *Bilder verifizieren in 5 Schritten* [How to Verify Pictures in Five Steps]. Retrieved from Netzwerk Recherche website https://netzwerkrecherche.org/wp-content/uploads/2017/06/Tipsheet-Bilder-verifizieren-nr17.pdf.

Gailītis, V. (2013, June 4). *Interview with Paul Hansen* [Web log post]. Retrieved from http://fkmagazine.lv/2013/06/04/interview-with-paul-hansen/.

Godulla, A. (2009). *Fokus World Press Photo. Eine Längsschnittstudie "ausgezeichneter" Pressefotografie von 1955–2009* [Focus World Press Photo. A Longitudinal Study of Exceptional Photography from 1955–2009]. http://d-nb.info/997407689/34.

Godulla, A. (2014). Authentizität als Prämisse. Moralisch legitimiertes Handeln in der Pressefotografie [Authenticity as a Premise? Morally Legitimized Actions in Press Photography]. *Communicatio Socialis, 47*(4), 402–410.

Wolf, C., & Godulla, A. (2016). Potentials of Digital Longforms in Journalism. A Survey Among Mobile Internet Users About the Relevance of Online Devices, Internet-Specific Qualities, and Modes of Payment. *The Journal of Media Business Studies, 13*(4), 199–221. https://doi.org/10.1080/16522354.2016.1184922.

Check for updates

CHAPTER 17

Photo Manipulation: Software to Unmask Tampering

Alessandro Tanasi

In the last 10 years, we noticed an unprecedented growth of multimedia files usage, digital images and videos, shared across social networks and other kind of communication technologies.

The access to more powerful and cheaper tools brought digital images editing to everyone and, with easy-to-use software for computers and mobile, made fraudulent image forgeries easier than ever.

The challenge of evaluating the trustworthiness of a digital photo has become a common demand and a booming research area.

To detect digital image forgeries and rate an image's fidelity, the science of digital image forensics comes to the rescue. In the last decades researchers have developed many digital image forensic techniques and tools.

A number of methods have been developed to detect image forgery; however, the process of verifying an image is a laborious and a time-consuming task. For example, the "Visual Verification Guide for Photos", provided by First Draft News (First Draft, n.d.), requires several steps involving human interactions.

A. Tanasi (✉)
Ghiro Project, Milan, Italy

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_17

Most techniques are based on the usage of footprints left on images. These artefacts produced and left could help to reverse-engineer and determine the history and authenticity of images.

Each device or process involved in image creation and manipulation usually leaves some kind of trace that could provide clues about the file. For example, camera and sensor noise, resizing, rotation, or filters leave artefacts which are important evidences of image manipulation.

While tampering techniques have become more sophisticated, the state-of-the-art image forensic tools and research in detection technologies against image forgery also keep up in an evolving and more sophisticated arms race.

Can image forensics always detect a forgery? It depends on who is better at their job: the investigator or the person who modified the image.

For the purpose of this chapter we are referencing only free tools—there are many great commercial image forensics products, but they are not affordable to all. For the following examples you are going to need only a Linux distribution like Ubuntu (Ubuntu, n.d.) or Kali (Offensive, Security n.d.) to run the software described.

First of all, it is important to understand how digital images are created: in most cases they are built using a graphics editor or captured with a digital photo camera.

In digital cameras, the light goes through an optical system, using a lens with a variable diaphragm to focus light, and falls on the sensor when the shutter is open. The sensor consists of a large number of photodiodes which act as light receptors and turns light into discrete signals.

Each receptor, dependent on the light intensity and colour, sends a different signal to the processor.

It translates signals in a digital image, and saves it on the memory, either in uncompressed format (RAW) or as compressed file (for example JPG). The file is created using a defined format, usually applying algorithms to encode data.

If the image is created or manipulated with a photo editor, the encoded data are decoded, processed as the user requested, encoded again, and saved.

Data are saved to disk in one of the many image file types, standardized formats for storing digital images and related data.

An image file may store data in compressed, uncompressed, or in vector format. A grid of pixels is composed by a number of bits; the size of the file is correlated with the resolution and the colour depth.

Images can be compressed with an algorithm to store an exact (lossless compression) or approximated (lossy compression) image representation.

There are literally hundreds of image file types, some standards are open and some proprietary.

The most often used image formats on the Internet are briefly described below:

- JPEG (Joint Photographic Experts Group) (JPEG, n.d.): Almost every digital camera can save photos in the JPEG format, the file extension used is .jpg or .jpeg. JPEG applies a lossy compression algorithm, which results in a huge file size reduction. However, when repeatedly edited and saved, the image becomes more and more degraded.
- GIF (Graphics Interchange Format) (World Wide Web Consortium 1987): It is a simple and aged format, created for storing graphics with few colours and using a lossy compression algorithm. Due to its animation capabilities, it is still used today. The file extension is .gif.
- PNG (Portable Network Graphics) (Duce 2003): free and open source format created as a successor of GIF. It excels with large images and uses a lossless compression algorithm. It is designed for online use and can be fully streamed with a progressive display option; the file extension is .png.

If you need to understand the format of a file, you can use the Linux "file" command ("file (command)" 2017): It looks to the header of the file to identify it against a database of known headers:

```
$ file unknown_file.xxx
unknown: PNG image data, 1534 x 960, 8-bit/color RGBA,
non-interlace
```

Other tools to identify format for unknown files are TrID (Pontello 2018) and Apache Tika (The Apache Software Foundation, n.d.).

If you are dealing with a large number of images, you could be interested in quickly identify exact duplicates or search for a specific image.

Image hashing is used to generate a hash value for an image. This hash value is a unique identifier for the image, and it can be used to search for that file, regardless of file name or extension.

Hash values can be used for image database indexing and for verifying content integrity.

Many hashing algorithms can be used, as MD5 or SHA256; for example, you can calculate SHA1 hash for a photo as follows:

```
$ sha1sum photo.jpg
da39a3ee5e6b4b0e3255bfef15601892afd90709 photo.jpg
```

The main issue with classic hashing is that one minor change to a file, and the hash (for example MD5) is completely different. A minor change, which leaves the photo the same to human eyes, is an alteration leading to a different hash.

Hash comparisons are either a yes or a no—either the hash matches, or it does not. But that does not mean that the files are not the same, it just means they are not exactly the same (in a bit-to-bit comparison).

A *fuzzy* or *perceptual hash* is a fingerprint of a file derived from various features from its content. Unlike cryptographic hash functions, it can efficiently and effectively help to identify images that contain a high percentage of similarities. An investigator can then take the files with the highest percentage of similarities and manually review those individual images.

SSDeep (Ssdeep Project 2017) and pHash (pHash, n.d.) are examples of fuzzy hashes. We are going to calculate SSDeep for an image, slightly modify it, and then compare it with the calculated hash, as follows:

```
$ ssdeep -b image.jpg
ssdeep,1.0-blocksize:hash:hash,filename
384:HEOV6N0/xFXSw0x2K+PFfNDOPK2TYWImaMsYLB3q60tL5DwpXe9hZ4ks-
JWaTNpyY:HEI9Xg7+P9yImaNk3qrDwpXe9gf5xkIZ,"image.jpg"
```

```
$ ssdeep -bm dez.hash altered.jpg
altered.jpg matches image.jpg (99)
```

The similarity between the two images is rated with a similarity of 99%.

Microsoft developed a technology dubbed PhotoDNA (Microsoft, n.d.) that computes the hash of multimedia files to identify images in order to prevent child pornography. Hashes for illegal images are computed and shared with trusted parties to search, block and, report these contents.

When a digital camera captures an image, it usually appends a whole set of additional information called metadata, which is stored within a file.

This data includes basic photo settings, as well as information about the device.

Photo editing software often adds its metadata including information about the manipulation when saving a file.

In a computer forensic investigation metadata can be extremely useful in answering some of the questions, such as when a photo was taken, or which camera was used.

Metadata are in most cases text information stored in the image following one or usually more than one Photographic Metadata Standards. Those have been developed by a number of organisations, the most common are:

- Exif (Exchangeable image file format) (CIPA 2012)
- XMP (Extensible Metadata Platform) (Adobe, n.d.)
- IPTC (International Press Telecommunications Council) (IPTC, n.d.)

Stored information following these standards can be found in most of the image file types—JPEG, PNG, TIFF, and more as well as many RAW images—and can be sorted in the following categories:

- Image properties: resolution, dimensions, compression
- Date and time: Most digital cameras and photo editing tools will record one or more timestamps

- Camera settings during capture: exposure time, shutter speed, focal length, ISO, camera orientation, if flash was used
- Camera hardware: model, serial number, lens information
- Thumbnail: A preview of the original image in small size is stored inside the metadata. It is used by file managers and camera's LCD screen when showing image's preview
- Descriptive text: keywords, description, copyright information

Some metadata tags can be a turning point in a forensic investigation. For example when the image is manipulated with Adobe Photoshop, it usually writes some tags in the XMP section describing the changes made.

The amount and type of metadata tags written can be different between camera models and vendors, and vendors sometimes also write proprietary tags, like the tag "MakerNote", containing custom or proprietary data.

This peculiarity can be used to fingerprint the camera, associate a photo to a vendor or a camera used, just looking at the metadata tags map.

Since metadata can contain relevant footprints, it is suggested to examine all the tags and their contents—a correlation between metadata and other evidence could quickly lead to close the investigation.

There are many tools available to read metadata, and many online services like FotoForensics (Hacker Factor, n.d.) or ImageForensic (ImageForensic, n.d.) are easy to use and require only a web browser.

Command line tools are also available, for example you can use ExifTool (Harvey, n.d.) to extract metadata (important data are highlighted in bold):

```
$ exiftool /tmp/IMG_0003.jpg

ExifTool Version Number         : 10.55
File Name                       : IMG_0003.jpg
Directory                       : /tmp
File Size                       : 1845 kB
File Modification Date/Time     : 2017:10:26 19:45:25+02:00
File Access Date/Time           : 2017:10:26 19:45:25+02:00
File Inode Change Date/Time     : 2017:10:26 19:45:25+02:00
File Permissions                : rw-r--r--
File Type                       : JPEG
File Type Extension             : jpg
MIME Type                       : image/jpeg
JFIF Version                    : 1.01
Exif Byte Order                 : Big-endian (Motorola, MM)
Make                            : Apple
Camera Model Name               : iPhone 6s
X Resolution                    : 72
Y Resolution                    : 72
Resolution Unit                 : inches
Software                        : Photos 2.0
Modify Date                     : 2017:08:07 20:05:21
Exposure Time                   : 1/950
F Number                        : 2.2
Exposure Program                : Program AE
ISO                             : 25
Exif Version                    : 0221
Date/Time Original              : 2017:08:07 20:05:21
Create Date                     : 2017:08:07 20:05:21
Components Configuration        : Y, Cb, Cr, -
Shutter Speed Value             : 1/950
Aperture Value                  : 2.2
Brightness Value                : 10.30948678
Exposure Compensation           : 0
Metering Mode                   : Multi-segment
Flash                           : Off, Did not fire
Focal Length                    : 4.2 mm
Subject Area                    : 2015 1511 2217 1330
Run Time Flags                  : Valid
Run Time Value                  : 5483907039916
Run Time Epoch                  : 0
Run Time Scale                  : 1000000000
Acceleration Vector             : -0.9775595067 0.03461138099 0.2647836766
Sub Sec Time Original           : 820
Sub Sec Time Digitized          : 820
Flashpix Version                : 0100
Color Space                     : sRGB
Exif Image Width                : 4032
Exif Image Height               : 3024
Sensing Method                  : One-chip color area
Scene Type                      : Directly photographed
Exposure Mode                   : Auto
White Balance                   : Auto
Focal Length In 35mm Format     : 29 mm
Scene Capture Type              : Standard
Lens Info                       : 4.15mm f/2.2
Lens Make                       : Apple
Lens Model                      : iPhone 6s back camera 4.15mm f/2.2
XMP Toolkit                     : XMP Core 5.4.0
Creator Tool                    : Photos 2.0
Current IPTC Digest             : 51b95a7af83cc54e5d1f6f373feab107
Coded Character Set             : UTF8
Application Record Version      : 2
Digital Creation Time           : 20:05:21
Digital Creation Date           : 2017:08:07
Date Created                    : 2017:08:07
Time Created                    : 20:05:21
IPTC Digest                     : 51b95a7af83cc54e5d1f6f373feab107
Image Width                     : 4032
Image Height                    : 3024
Encoding Process                : Baseline DCT, Huffman coding
Bits Per Sample                 : 8
Color Components                : 3
Y Cb Cr Sub Sampling            : YCbCr4:2:0 (2 2)
Aperture                        : 2.2
Date/Time Created               : 2017:08:07 20:05:21
Digital Creation Date/Time      : 2017:08:07 20:05:21
Image Size                      : 4032x3024
Megapixels                      : 12.2
Run Time Since Power Up         : 1:31:23
Scale Factor To 35 mm Equivalent: 7.0
Shutter Speed                   : 1/950
Create Date                     : 2017:08:07 20:05:21.820
Date/Time Original              : 2017:08:07 20:05:21.820
Circle Of Confusion             : 0.004 mm
Field Of View                   : 63.7 deg
Focal Length                    : 4.2 mm (35 mm equivalent: 29.0 mm)
Hyperfocal Distance             : 1.82 m
Light Value                     : 14.2
```

Thanks to metadata we can see that the photo was taken with an iPhone, powered up for one and a half hours, and that, when saved, the file was downloaded using Photo app.

For performance reasons, cameras store a preview inside metadata; this thumbnail can be extracted from one file to another with ExifTool, with the following command (with img_003.jpg being the original image and preview.jpg the extracted thumbnail):

```
$ exiftool -b -PreviewImage /tmp/img_003.jpg > /tmp/preview.
jpg
```

It is interesting how in rare cases, if the image is edited with a software that is able to save alteration to the image but misses to update the preview inside metadata, by extracting the preview you can spot the difference and discover the original image before manipulation.

The Exif format has standard tags for location information, and most of the smartphones have a built-in GPS receiver enabled during capture to include the location information in the metadata.

Geotagging is the process of adding geographical identification metadata to media; this data usually consists of latitude and longitude coordinates and the altitude.

The location is calculated using GPS, WiFi router locations, or cell tower identification preferring the most accurate method available.

When stored in metadata, latitude and longitude are stored in units of degrees with decimals. The location values can be read by many programs, such as ExifTool:

```
$ exiftool -location:all /tmp/test.jpg
File not found: exiftool
======== /tmp/test.jpg
GPS Latitude Ref                : North
GPS Longitude Ref               : East
GPS Altitude Ref                : Above Sea Level
GPS Img Direction Ref           : True North
GPS Img Direction               : 82.12307692
GPS Altitude                    : 0 m Above Sea Level
GPS Latitude                    : 38 deg 54' 35.40" N
GPS Longitude                   : 1 deg 26' 19.20" E
GPS Position                    : 38 deg 54' 35.40" N, 1 deg 26' 19.20" E
```

If further clues about the geolocation of the photo are needed, what you see can be visually compared with other images that can be found online.

Tools as Google Street View (Google, n.d.-a) offer valuable location intelligence, as it can provide a virtually 360-degree view of any location including buildings and landscaping.

Since metadata are really easy to read, write and delete, it is not possible to completely rely on them as they can be easily modified.

Images are a digital translation of a picture, using algorithms colours and light written as bits. When an image is edited, the original mathematical properties are changed and could reveal anomalies.

The tampering localisation algorithms were created to detect the most common types of tampering footprints that can be detected in an image, just by looking at its properties.

Error Level Analysis (ELA) ("Error level analysis" 2017) is one of the dominant image tampering localisation methods currently used. It is based on the idea that each re-saving of a JPEG file compresses it with lower quality. An area that was edited would have a different compression level, so under ELA it would be brighter than other areas.

Other methods are established as standards in major image forensics platforms: Discrete Wavelet High Frequency Noise, Double JPEG Quantization, Median Filtering Noise Residue, JPEG Blocking Artefact Inconsistencies and JPEG Ghosts.

Each technique aims at detecting different tampering traces; there is no silver bullet, and not all algorithms should be expected to work on all images. Therefore, it is suggested to use different algorithms and correlate the traces found.

Light and shadows are an inseparable factor of an image. There are several techniques to detect manipulated images based on light source direction, light intensity, or shadow.

The brightness of a certain region of an image varies with the variations in the shape of the real object. One or many inconsistencies in the lighting environment or light source direction on various parts of the image can be used to detect tampered images.

An image can be manipulated to add more information (e.g., more trees in a wood) or to remove some of it (e.g., fewer trees in a forest, copy-moving the grass over real trees), using the *image splicing* or *copy-moving* manipulation. The former refers to the practice of copying a part of an image and inserting it into another, so as to give the impression that an additional element was present in a scene. The latter means taking a part of an image and duplicating it in another area within the same image.

These two practices have an important distinction as different techniques can be used to identify each one.

The detection of copy-move forgeries is based on finding internal replications of parts or blocks of the image. Splicing detection is more complex and is usually based on the assumption that the spliced region differs from the rest of the image in some significant aspect.

The most frequent misuse of image duplication is ripping images, the theft of intellectual properties. A person who intends to copy images without permission creates a near duplication of the original images.

This is a growing phenomenon in eyewitness journalism, for example in breaking news: In case of a terrorist attack, Images from previous attacks are published on social media, duplicated or altered to seem up-to-date.

In addition to image forgery detection, the detection of duplicates and near-duplicate images is emerging as an important forensic challenge.

A variety of existing online services such as Google Images (Google, n.d.-b) or TinEye (TinEye, n.d.) can aid in the process of duplicate detection and social media content verification.

Their use is suggested as a best practice, to quickly check the uniqueness of an image or to search for similarities.

Sometimes images can be tampered in an extremely professional way, using specific techniques to avoid tampering detection called *anti-forensics.*

Though many existing forensic techniques are capable of detecting a variety of common image forgeries, usually they do not account for the possibility that anti-forensic and ad-hoc custom operations may be designed and used to hide manipulation footprints.

Anti-forensics require skilled professionals and expensive services; in most cases, you can ignore the probability to deal with images of this kind, except if your threat model is designed for high profile, well financed, and skilled actors.

We live in a time where multimedia and digital images distributed over social networks have widespread impact on people and societies.

The presence of metadata in an image is by no means guaranteed. In fact, for several years, it has been observed that most social media platforms tend to strip metadata from uploaded multimedia to protect their users' privacy. Sometimes images from social networks or messaging application are also resized or cropped.

Investigation on images coming from social media should take care of this sanitisation and focus on forensic techniques not based on metadata.

Researchers have developed image forensics tools and frameworks, some of them being on-premises tools, like Ghiro (Ghiro, n.d.), Jpegsnoop (ImpulseAdventure 2017), or Phoenix (Fırat, n.d.), and others being online services, like ImageForensic or Reveal (Reveal, n.d.).

If you are evaluating an image forensic tool it is suggested it should meet at least the following requirements:

- Easy integration in your image verification workflows
- Easy to use, also for non-skilled people with little training
- Graphical interface and visual analysis
- Summary and exhaustive results for a quick overview and deep analysis; clear indicators describing key image manipulations
- Support of metadata extraction, image search, duplicate detection, image classification
- Access to the source code or at least an explanation of the techniques used under the hood. Never trust "black box" tools, as an investigator you should understand the methods you are using.
- Updated with the latest advances in image forensics techniques

Image forensics is not a matter of tools only, it is also about processes and people. Workflows should take care of image validation, and people should be trained to understand the problem of image authenticity and to know how much trust to put in digital media.

References

Adobe. (n.d.). *Extensible Metadata Platform (XMP)*. Retrieved from https://www.adobe.com/products/xmp.html.

CIPA. (2012). *Exchangeable Image File Format for Digital Still Cameras: Exif Version 2.3*. Retrieved from http://www.cipa.jp/std/documents/e/DC-008-2012_E.pdf.

Duce, D. (Ed.). (2003). *Portable Network Graphics (PNG) Specification* (2nd ed.). Retrieved from http://www.libpng.org/pub/png/spec/iso/index-object.html.

Error Level Analysis. (n.d.). *In Wikipedia*. Retrieved 2017, from https://en.wikipedia.org/wiki/Error_level_analysis.

file (command). (n.d.). *In Wikipedia*. Retrieved 2017, from https://en.wikipedia.org/wiki/File_(command).
Fırat, B. (n.d.). *Phoenix—Image Forensics*. Retrieved from https://github.com/ebemunk/phoenix.
First Draft. (n.d.). *Visual Verification Guide—Photos*. Retrieved from https://firstdraftnews.org/en/education/curriculum-resource/visual-verification-guide-photos/.
Ghiro. (n.d.). [Homepage]. Retrieved from http://www.getghiro.org/.
Google. (n.d.-a). *Google Street View*. Retrieved from https://www.google.com/streetview/.
Google. (n.d.-b). *Google Images*. Retrieved from https://images.google.com/.
Hacker Factor. (n.d.). *FotoForensics*. Retrieved from https://fotoforensics.com.
Harvey, P. (n.d.). *ExifTool by Phil Harvey*. Retrieved from https://sno.phy.queensu.ca/~phil/exiftool/.
ImageForensic. (n.d.). [Homepage]. Retrieved from https://www.imageforensic.org/.
ImpulseAdventure. (2017). *JPEGsnoop 1.8.0—JPEG File Decoding Utility*. Retrieved from http://www.impulseadventure.com/photo/jpeg-snoop.html.
IPTC. (n.d.). *Photo Metadata*. Retrieved from https://iptc.org/standards/photo-metadata/.
JPEG. (n.d.). [Homepage]. Retrieved from https://jpeg.org.
Microsoft. (n.d.). *Photo DNA Cloud Service*. Retrieved from https://www.microsoft.com/en-us/photodna.
Offensive Security. (n.d.). *Kali Linux*. Retrieved from https://www.kali.org.
pHash. (n.d.). pHash. *The Open Source Perceptual Hash Library*. Retrieved from http://phash.org/.
Pontello, M. (2018, April 8). *TrID—File Identifier*. Retrieved from http://mark0.net/soft-trid-e.html.
Reveal. (n.d.). *Reveal Image Verification Assistant*. Retrieved from http://reveal-mklab.iti.gr/reveal/.
Ssdeep Project. (2017, November 7). *Ssdeep—Fuzzy hashing Program*. Retrieved from http://ssdeep.sourceforge.net.
The Apache Software Foundation. (n.d.). *Apache Tika—a Content Analysis Toolkit*. Retrieved from https://tika.apache.org/.
TinEye. (n.d.). *Reverse Image Search*. Retrieved from https://www.tineye.com/.
Ubuntu. (n.d.). [Homepage]. Retrieved from https://www.ubuntu.com/.
World Wide Web Consortium. (1987). GIF. Graphics Interchange Format. *A Standard Defining a Mechanism for the Storage and Transmission of Raster-Based Graphics Information*. Retrieved from https://www.w3.org/Graphics/GIF/spec-gif87.txt.

PART III

Authenticity, Identity and Transparency

The previous chapters have shown emerging techniques that challenge the status quo of journalism every day. A continuous datafication of sources and information flows and accompanying requirements challenge not only practitioners, but in more general journalism as a pillar of society. Innovative approaches, more than ever relying on technological innovations, prompt producers and editors to incorporate new practices into their repertoires; moreover, not only journalists, but also their readers and users who are a manifestation of the public sphere need to catch up with these innovations.

Therefore, it is imperative to discuss previously introduced techniques and projects from a societal and sociological viewpoint. Adducing current incidents, the following chapters examine aspects and examples that are both results from an ongoing digitalisation and evolution of journalism on the one hand, and on the other hand reasons for the adaptation of innovative approaches. The dilemma here is a constantly widening gap between technological advancements and deadlocked regulatory bodies and policies.

Authenticity, identity and transparency challenge journalism and other societal institutions on several levels. Information warfare and propaganda try to undermine and destabilise established systems with the objective to influence public opinion and to manipulate marginal groups. As the political sector and the public sphere are evidently prone to attacks from maleficent influencers, the journalistic system and its logic—serving as a corrective in public communication and opinion building—need to be discussed against this backdrop.

Check for updates

CHAPTER 18

Fact-Checking as Defence Against Propaganda in the Digital Age

Anna Sarmina

Globalisation, which covered all areas of human activity and became a component of global integration processes, has led primarily to significant changes in the information space. The result of global changes in almost all areas of human life and activity involved a transformation of society, therefore information gained a radically new meaning. Information now has a global character; it means that the information flows cannot be stopped by state borders or other barriers. Digitisation promoted new ways for gathering, storing, and disseminating information, as a result information has become an important and even strategic resource that is produced and consumed in society. Furthermore, the information space causes social changes.

All phenomena associated with the information belong to the field of information policy that deals with the organisation of the processes of origin, dissemination and storage of the information in social systems. Information reflects reality, and having control of information facilitates the control of reality, in particular because each state strives to control as many information streams as possible (either officially or unofficially). This caused the creation of different forms of influence on mass

A. Sarmina (✉)
Taras Shevchenko National University of Kyiv, Kiev, Ukraine

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_18

consciousness, for instance think tanks—information interventions that aspire to transform the public information space by using own resources (economic, political, military, and international).

In this context it is emphasised that the information space changes all the time, producing ever new instruments for influencing the public. The more invisible such instruments are, the more powerful they are. First of all, propaganda is one of those instruments. Propaganda can be successful only if it is inconspicuous, in which case it has a destabilising effect. Thus, since the role of the information space is really important, it is to be kept and defended, which is one of the main tasks of the information policy of each state.

The radical realisation of information policy takes the form of an information war. This means that information mechanisms are used to solve non-informative tasks. As a result, non-informative objectives are achieved through the use of information tools, and thus the mass consciousness is transformed. It is to be underlined that the objectives of modern information wars cover a wide range of areas, for example economics, politics, culture, military, and other spheres.

Information war is to be seen as intervention in the information space of another state. Such intervention can pursue two main objectives—to integrate information or to steal information. So, the information that belongs to the opponent must either be destroyed or controlled, while own information must be protected. Information infrastructure, therefore, has already become an official object of attack and defence. The Prussian theorist Carl von Clausewitz maintained in his work "On war" that "war is a mere continuation of politics by other means" (von Clausewitz 2010). Generally, "by other means" describes armed violence that is nowadays understood in the form of combat. Four generations of wars are known in the newer history, which began after the Westphalian peace. At the moment, we are experiencing the fourth and the most modern generation of war: information warfare.

But what is of decisive importance for this kind of war is the fact that the information warfare is no longer a complementary but a decisive one. The clash of information and disinformation, various types of counterfeits (fakes)—all this is undoubtedly part of the information war. Furthermore, the crucial role of information in the war of the fourth generation is that the information thereby gives the possibility to conquer without armed struggle. This also implies the tendency to distance oneself from the armed struggle; many states, as well as Ukraine, have

become demilitarised zones. Therefore, there is a need to use special weapons, mainly cyber weaponry and information weaponry. One of the main features of the 4th generation war is mediatisation because this war is running under the eye of the camera and transmitted in media. Marshall McLuhan defined this type of war very accurately. He described it as the third war and claimed that "World War III is a guerrilla information war with no division between military and civil participation" (McLuhan 2015).

It is clear that information campaigns are primarily targeted at informative resources, but they nevertheless produce physical consequences. The modern type of information war is the so-called "hybrid war" or "non-linear war". Hybrid war is a revolution in the form of terrorism. It is the mixture of subversive activity, ideological influences, informational influences, and ordinary military operations. It is therefore a society-specific warfare, in which irregular and regular tactics are combined with modern information.

The goal of this hybrid war is to change the character of power and to influence the opinion of the civilian population of the opponent. It is the war of external influences and it includes a wide range of different warfare operations, including conventional and irregular combat, terrorist attacks, information operations, or espionage. Information plays a significant role in hybrid wars because its main task is to camouflage physical violence. Glenn Russel clarified this phenomenon by talking about a simultaneous combination of political, military, social, and media means with conventional and irregular methods of terrorism and of subversive-criminal warfare (Russel 2009).

Thus, in every type of information war all possible means are used as weapons, in addition to conventional methods. An important part of any information war is information weaponry, which distinguishes it from other wars. Information weaponry comprises complex technologies of destructive influence on information resources and information systems of the opponent, on the consciousness and on the area of the unconscious of the opponent. The information weapon is universal, accessible, and massive, has non-lethal character, destroys network systems (with viruses), controls information resources, and influences radically.

The use of information weapons in wars is closely linked to the art of propaganda. Propaganda is as old as the hills. The origin of the word is neutral: It comes from the ecclesiastical sphere and used to be an honourable word. The term has received negative shading only later,

primarily because it was compromised by political regimes. Under propaganda one understands intensive communicative processes which are aimed at changing the behaviour of the audience. Theoretically, propaganda is a special form of systematically planned mass communication, which does not want to inform or argue, but wants to persuade or convince (Bussemer 2015). It intends to persuade to take a specific position on a specific issue and to act according to that conviction.

Disinformation is also an influential weapon of the information war. It is the targeted distribution of false or misleading information, therefore it is to be seen as fraud or one of its forms—lie, deception and others. Disinformation and propaganda have much in common—both are tampering with information, misleading people by producing and presenting untrue information, creating fakes and spreading them mainly by mass media. In contrast to propaganda, which appeals to "emotion", disinformation manipulates with "ratio". But the main aim of both techniques is that truths should be distorted, lose their value, and become useless. That is, the first strangled victim of the information war is the truth.

The modern period is indicated as "Post-Truth Era", which could be explained as "relating to or denoting circumstances in which objective facts are less influential in shaping public opinion that appeals to emotion and personal belief" (Oxford English Dictionary). It means that the contributed information is inaccurate, skewed, biased, or fabricated. This Post-Truth Era generates some new phenomena, for example "fake news". Fake news are distorted signals uncorrelated with the truth (Allcott and Gentzkow 2017). Reporting mistakes, rumours, conspiracy theories, satire, false statements, or slanted reports are disseminated with special intentions (Allcott and Gentzkow 2017).

Fake news, propaganda and disinformation are powerful weapons in the modern world, as they play a significant role during the information war, and all over the world they are changing the media landscape step by step. The development of social media has not only made a great contribution to global changes, it has also affected democracy. Social media is seen as a special mechanism of disseminating fake news as the information can be posted without prior approval and an increasing number of users are getting their news from social media. The spreading of fake information through social media is the best way because most of the social groups accept information without checking or verifying it. In this way society pollutes the information environment by itself. A lot of people are sceptical of the information spread by state, but they express a

greater degree of confidence in what their friends are sharing. Actually, fictitious messages such as fake news exploit the functional logic of social networks.

At first blush, it is difficult to realise the gravity of consequences that fake news are able to cause. But a deeper approach to this problem makes it clear that fakes can be really dangerous insofar as they can damage the reputation of a person as fake producer as well as fake consumer. Moreover, fakes create distrust in all types of media, and in this way they damage democracy. False information can ruin lives and politics, and construct new reality by threatening through lie. The aim of spreading fakes is to create chaos in civil society. A wide range of examples of the negative consequences of fake news is well known; among them "Pizzagate" or the US election result in 2016 are worth mentioning.

Various types of fake news are known from media sources, for instance "100% False", "Slanted and Biased", "Pure Propaganda", "Misusing the Data", "Imprecise and Sloppy", "Misinformation", and "Shitposting" (Johnson 2016). The verification agency First Draft News describes seven types of Mis- and Disinformation. Among them are "False Connection", "False Context", "Manipulated Content", "Satire or Parody", "Misleading Content", "Imposter Content", and "Fabricated Content" (Wardle 2017). Nevertheless, propaganda still has the leading role.

Nowadays, at the time of Russian interventions in democratic processes in Ukraine, reliable sources are significant. The actual events in Ukraine are a mixture of propaganda, subversive activity and, of course, military actions. So the Ukraine conflict is the obvious pattern of a hybrid war that consists of two inalienable components—information attacks and military actions. Russia's policy of destabilisation in Ukraine has attracted the attention of experts to the phenomenon of the information war and its tools. "Russia had to make a big propagandistic effort – to change the situation in the physical domain, it had to change it in information and virtual domains" (Pocheptsov 2017).

In Russia, information campaigns are closely linked to propaganda—perhaps because Russia has inherited from the USSR a certain experience of using propaganda for the purpose of psychological influence on individuals. The Cold War was based on propaganda mechanisms. The information war between Ukraine and Russia has a whole series of examples of "twisting facts" and manipulation with information. The annexation of the Crimea was accompanied by the massive use of propaganda and

disinformation being spread through different channels: by TV broadcasters, by radio, by newspapers, through the Internet. The Maidan revolution was portrayed as fascist and extremely russophobic. According to the main types of propaganda (Lee and Lee 1979) all of them were used as weapon in Ukraine conflict. For example, during the annexation of the peninsula, explicit names for Russian soldiers were avoided in Russian reports and instead replaced by such euphemisms as "little green men" or "polite people". This is a brilliant example of the *card stacking* technique.

The *glittering generalities* technique was aimed at the destruction of characteristics of illegality; negative labels were neutralised by using positively connoted components (for example "people's governor") and by replacing illegal components with positive concepts, for example "reunification of Crimea" instead of "annexation" or "self-defence" instead of "occupation".

The *name calling* technique was presented by the reinforcement of negative characteristics of the opponent. By introducing the Maidan revolution, the protesters were called "radicals", "fascists", "extremists", "nationalists", and "junta". In opposition to this, the typical name for pro-Russian elements was "patriots".

As a brilliant example of the *plain folks* technique could be considered the (false) information that a woman assured she witnessed the Ukrainian soldiers crucifying a small boy and then tying his mother to tank and dragging her.

The utmost importance has the fact that during the annexation of the Crimea, telephone and Internet connections were cut by intentional cables damage. In order to create a media vacuum, television was taken under control, the Ukrainian channels were taken off the air, and Russian programmes were broadcast. So this approach came with a special focus on broadcast media.

Fakes have become crucial elements of this hybrid war. Aggressive propaganda activity is aimed at domination in the information sector. In order to achieve this goal, all possible means are applied—activity of information services, special services, international terrorist groups. All of them strive for information dominance, by using terror, propaganda, diversion, and sabotage holding attacks on information, on rights, and on liberties of the state as well as of humans. These methods demoralise society, and lead to confrontation in the society and to anarchy as a result.

As a matter of fact, fighting propaganda is a real challenge. Certainly, in order to defend yourself, it is important to become critical and to prevent negative consequences of fake news. It is necessary to debunk and disable them.

There is a wide range of special organisations that keep debunking false information. Nevertheless, common Internet users are included to do debunking by comments (first of all in social networks).

Fact-checking is the most popular method of debunking, and in this way it is an important tool for protecting the truth. Fact-checking is relevant in all spheres of everyday life, but primarily it is important for journalism, so that people could trust media. It helps to keep the media climate healthy. "If journalism is a cornerstone of democracy, then fact-checking is its building inspector, ensuring that the structure of a piece of writing is sound" (Borel 2016). Journalism plays a crucial role in the process of formation of democracy. Moreover, it is so important to maintain public trust and to become an anchor of truth in the ocean of propaganda. We became witnesses of a new Digital Era that causes global changes across the world and has changed the role of information. Earlier, the message was important, now interpretation, which is the reason why information wars evolve. Due to the digitisation, the boundaries of what can be said and perceived have been moved. Furthermore, digitisation facilitated the creation as well as the debunking of fakes.

The aim of fact-checking as a democratic institution is promoting truth in public discourse. Primarily it has become the task of NGOs, think tanks, civil society, and other organisations to counter blatant propaganda. During the last years a significant rise of fact-checking websites has been fixed. The statistics say that nowadays there are 104 active websites (the total number is 169 according to Duke Lab) whose activity is aimed at debunking fakes. To compare, in 2014 there were only 44 web pages. This indicates the growing need for debunking fakes, and at the same time it is the signal of an increasing number of fakes. Almost every country of the world has at its disposal at least one web resource. For example FactCheck.org in the USA, Faktomat in Germany, Les Decodeurs in France, Factcheckni.org in Great Britain, Pagella Politica in Italy, Faktabaari in Finland, Factchecker in India, the Conversation Fact Check in Australia, Aos Fatos in Brazil, and Trudeau Meter in Canada.

In Ukraine there are at least four active web resources that evaluate the truth in news reports. The most reputed is StopFake that was founded by alumni and volunteers at the Mohyla Journalism School in

Kiev. This website was created in 2014 for debunking Russian propaganda in Ukraine and during the last years it has grown spectacularly. Now, this resource provides information in ten languages not only about debunked fakes, but it also serves as an educational resource for civil society. Ukraine is still living under fake news bombardment that contributes to the manipulation of public opinion. In our society the media informs people primarily. As a rule, this way of informing involves huge audience. So StopFake was launched to check facts, verify information, refute false news about Ukraine, and this challenge was faced first of all to restore public trust in the media. Furthermore, they keep protecting journalism standards and draw public attention to these standards.

They analyse thousands of reports about Ukraine and in case they establish the fact of a submission of fake information they start to investigate it. Herewith they check the sources of fakes and give conclusions about unreliable information. It is not always possible to trace the tracks of the origin of rumours. Nevertheless, the strategy becomes apparent.

Actually, the common techniques and methods for debunking are used, for example checking the source (or even double-checking), checking geolocation, and verification of the information. In some cases it is possible to investigate the origin of the rumour. In order to debunk fakes, the possibilities of Google are widely used, as well as tools for checking pictures in order to authenticate user-generated photos. The most common methods are verifying information about a website and its owner, geolocation verification, using metadata, and asking witnesses.

In this way StopFake as well as other organisations whose activity is devoted to debunking has become a real power in an information war, being able to disprove fakes. StopFake is a first line defence against propaganda in Ukraine.

For example, at the beginning of May 2017 Russian media reported about a peak of tourism from Ukraine in annexed Crimea. But later on, the Border Service Advisor of Ukraine refuted this information. On 2 May 2017 it was reported that after a meeting with Russian President Vladimir Putin, German Chancellor Angela Merkel supports the suspension of sanctions against Russia. However, the statement of Merkel was reduced and quoted out of the context, and in this way "twisted". The Chancellor noted that sanctions would be lifted only after the elimination of the reasons for which they were imposed.

A wide discussion was prompted by the topic of a beginning famine in Ukraine. Russian media published a post from Facebook with a photo

and comment that Ukrainians are on the verge of starvation. Other media added the false information about introduction of food cards in Ukraine. This story was based on a video about the introduction of a programme similar to the US food stamp programme in order to control the prices of some food.

The fact that fighting against fake news is a top priority for Ukraine is confirmed by the premiere of the documentary film "Nothing but Lies: Fighting Fake News" (by Tim White), presented in Kiev in March 2017. The film shows the history of fake news as well as actual methods and ways for spreading them in the societies.

In order to draw conclusions, it is necessary to point out that the Digital Age has brought significant changes that have spread over all spheres of life. Primarily, these changes have an impact on the information domain. Globalisation has brought in addition to a number of advantages new ways for the inflow of foreign information products, which propagate foreign values. Due to the development of new media, the prohibitions that existed earlier were removed; moreover, new opportunities for terrorism have appeared.

Since 2014, we have been witnessing in Ukraine the advanced form of hybrid warfare run according to atypical patterns that consists of armed conflict, violence and propaganda as invisible power in this war. Actually, non-military means play a more significant role than weapon. Nevertheless, the strategy has become apparent; the rising fact-checking and debunking activities have become the reaction on the propaganda machine and disinformation company. Independent fact-checkers are fighters in an information war.

Information campaigns are by no means a new phenomenon because information has always been a part of conflict; such campaigns have already evolved in the twenty-first century. They are characterised by the use of communication technologies as well as by cyber attacks, and by the growing importance of civil domains. In different countries they take different forms, and different tactics are used. Nowadays, we are among numerous information flows and that is why every information, every fact, every event must be subject to critical analysis.

References

Allcott, H., & Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives, 31*(2), 211–236. https://doi.org/10.1257/jep.31.2.211.

Borel, B. (2016). *The Chicago Guide to Fact-Checking*. Chicago: The University of Chicago Press.

Bussemer, T. (2015). *Propaganda: Konzepte und Theorien*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Duke Reporters Lab (Cartographer). (n.d.). Global Fact-Checking Sites. [Interactive Map]. Retrieved from https://reporterslab.org/fact-checking/.

Johnson, D. J. (2016, December 13). The Five Types of Fake News. *Huffington Post*. Retrieved May 15, 2017, from http://www.huffingtonpost.com/dr-john-johnson/the-five-types-of-fake-ne_b_13609562.html.

Lee, A. M., & Lee, E. B. (1979). *The Fine Art of Propaganda: Prepared for the Institute for Propaganda Analysis*. New York: Octagon Books.

McLuhan, M. (2015). *Culture Is Our Business*. Eugene, Oregon: Wipf & Stock Publishers.

Pocheptsov, G. (2017). The Origins of Fake and Alternative Facts Can Help Us Understand the Concept of Post-truth. *Russian Journal of Communication, 9*(2), 210–212. https://doi.org/10.1080/19409419.2017.1323180.

Post-Truth Era. (n.d.). *In Oxford English Dictionary*. Retrieved May 15, 2017, from http://www.oed.com/.

Russel, W. G. (2009). *Thoughts on "Hybrid" Conflict*. Retrieved from http://smallwarsjournal.com/blog/journal/docs-temp/188-glenn.pdf%3Fq%3D-mag/docs-temp/188-glenn.pdf.

von Clausewitz, C. (2010). *On War: Volume I* (n.p.). Floating Press.

Wardle, C. (2017, February 16). Fake News. It's Complicated. *First Draft*. Retrieved June 15, 2017, from, https://firstdraftnews.com/fake-news-complicated/.

Check for updates

CHAPTER 19

Crowdsourced and Patriotic Digital Forensics in the Ukrainian Conflict

Aric Toler

The Ukrainian Conflict is the first European war to be fought with the ubiquitous presence of the internet, producing thousands of hours of video and over three years' worth of online witness accounts relaying information about the ongoing war. While this wealth of materials allows analysis of the war in ways that would be a pipedream for journalists and historians a century ago, there is a gap between the amount of materials compared to the capabilities and available time for professional journalists to review them. While the Syrian Civil War has lasted for over two years longer than the Ukrainian Conflict, a similar problem exists there, where, according to Kosslyn and Green (2016), "there are more hours of footage online than there have been hours of actual conflict." Just as with the Syrian Civil War, there are years of footage online concerning the war—more than professional journalists alone are able to handle, especially with the recent downsizing of newsrooms.

From the start of the Ukrainian Conflict in 2014, this vacuum was filled by crowdsourcing, with amateur enthusiasts finding, verifying, and analyzing online materials related to the war. The individuals and groups that became active in this sphere had different motivations and expertise,

A. Toler (✉)
Bellingcat, Kansas City, MO, USA

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_19

with some choosing to maintain an anonymous internet persona and others who use developing methodologies and practices of crowdsourced online research to spur real-world action. But what makes these research groups different than other types of investigators? For one, there is a decentralized nature to them, with contributors flung across the world and no central newsroom, as you would see with most newspapers and even online publications. There is also the distinction of the types of materials they use to build investigations, which are openly available to anyone with an internet connection, as opposed to research conducted on the ground and through local sources. While some research collectives use some non-digital research in their investigations, the vast majority of the research is built off of open source materials, which bring a number of challenges in verification and contextualization.

What Materials Are Being Investigated?

Videos are often the most important element of open source analysis related to the Ukrainian Conflict. The sources of the majority of these videos are:

1. Professional media outlets, including both domestic (Hromadske, TSN, 112 Ukraine…) and foreign outlets (BBC, RT, CNN…). Also included are semi-professional media outlets that often veer into propaganda, such as News-Front and DONi News.
2. Governments, including information ministries. Some of these include Ukrainian Military TV and the so-called Donetsk People's Republic (DNR) Ministry of Information.
3. Participants of the conflict, including both enlisted soldiers from Ukraine, Russia, and separatist forces, and members of volunteer battalions.
4. Citizens, including those living in the conflict zone, and others who record footage relevant to the on-going war, such as military convoys en route to the Russia-Ukraine border or conflict zone.

We can see a nearly full range of these types of sources in tracking a single BMP-2, a Soviet amphibious infantry fighting vehicle, with the word Lavina ("Avalanche") written across it in Russian. In August 2014, a Russian woman named Anna Bocharova uploaded a photograph (Bocharova 2014) onto the Russian social network Vkontakte

of "Lavina," as it was being hauled through the Russian city of Staraya Stanitsa, near the Ukrainian border. She left a geotag on the photograph, making it easy to discover the location, along with other geographical details that can be found on satellite imagery. At about the same time, a Russian filmed (Action Tube 2014) as a convoy of military vehicles passed by in the same city. "Lavina" can be seen passing by, along with other vehicles, including one with the phrase "For the Donbas!" written on the side. This video was uploaded to YouTube, among other social networks.

This was not the last we would see of Lavina. A month after the Russian citizens filmed Lavina in their country, September 2014, an Associated Press reporter in Ukraine photographed (Vojinovic 2014) Lavina in the city of Zhdanivka in the Donetsk Oblast. In October, a Ukrainian citizen filmed a video (Roman iNapalm 2014) of Lavina travelling through separatist-controlled Yenakiieve, Ukraine. Later, a man named Sergey who fought alongside separatist forces in Donetsk posted a picture of himself on Vkontakte posing with Lavina ([Digital photo], n.d.), while holding a rifle. Finally, in 2015, Lavina reached Vuhlehirsk, where it participated in the February 2015 Battle of Debaltseve, which led directly to the second Minsk ceasefire agreement. Videos from the propaganda news outlet News-Front (2015) along with the major Russian news publication Komsomolskaya Pravda (Bellingcat Vehicle Tracking Project 2015) show Lavina in the winter battlefield where Russian and separatist forces drove Ukraine out of the key rail hub of Debaltseve.

Together, all of these photographs and videos provide us with a timeline of one Russian military vehicle travelling from a border city in Russia, to Ukraine, and finally to a major battle that tipped the scale of the Ukrainian Conflict. Most journalists do not have the time or resources to look for all of these sources—while some are easy to find, such as the battle footage from a major media outlet, others are not, such as the photographs shared by normal Russian and Ukrainian citizens, or participants of the conflict. However, a number of groups did have the collective time and expertise to gather all of these data points.

Open Source Research Groups

The three largest investigative groups that emerged during the Ukrainian Conflict are Bellingcat, InformNapalm, and the Conflict Intelligence Team (CIT). Each of these groups operates with different methodologies

and structures, but they are united in their near-simultaneous rises and their focus on the Ukrainian Conflict, and later with Russia's involvement in the Syrian Civil War. In particular, the groups differ in how political ideology is expressed in their research, their levels of cooperation with larger media outlets and independent groups, and their efforts to include reader engagement with research.

Conflict Intelligence Team (CIT)

The Russia-based CIT was founded under the name "War in Ukraine," and then rebranded to its current name in September 2015 after Russia began its military operations in Syria. Over the past two years, this group led by the Muscovite Ruslan Leviev has published a number of investigations that have garnered international recognition on the presence and activities of the Russian military in Ukraine and Syria. Notably, CIT has uncovered information regarding the deaths of Russian soldiers deployed to Syria and Ukraine, along with uncovering information regarding the Russian Armed Forces' use of cluster munitions in Syria. An example of one of their most successful investigations was researching photographs and video footage from Russian state-operated news outlets from the Russian air base in Syria's Latakia province. After intense analysis of these photographs, there was clear evidence of cluster munitions being deployed by the Russian Air Force in bombing runs throughout the country. Combined with photographs shared from Syrian civilians, this research showed—entirely through open source materials—that Russia was employing these munitions, despite denials from officials.

CIT is unique among open source-focused investigative groups in that they are a very small team, with only six members, and all of their researchers are Russian. The fact that this group is made up entirely of Russian citizens and primarily operates in Russia has brought them, and especially founder Leviev, unwanted attention from Russian authorities. Leviev was targeted in a hack perpetrated by the Kremlin-linked hacker group CyberBerkut (n.d.). Leviev has not just been targeted online, but he has also been harassed with "phone calls, death threats" due to the nature of his investigations (D'Agata 2016). Since he began probing the hidden activity of the Russian Armed Forces, Russian authorities have twice attempted to open a criminal case against him, including in an

instance where he was interrogated by a military prosecutor.[1] As of yet, the charges have not stuck, but the threat remains due to a 2015 decree that makes military deaths during peacetime a state secret (Luhn 2015).

Kirill Mikhailov, a Russian anti-Kremlin activist who now lives in Ukraine, joined the CIT in 2015 during an investigation into the death of three GRU agents in Luhansk, Ukraine. He joined in order to "uncover the truth about Putin's war in Ukraine," and does not hide that the team has a degree of political activism built into its investigations. Dating back before the annexation of Crimea and the war in the Donbas, Leviev has closely cooperated with Russian oppositionist and anti-corruption activist Aleksey Navalny. Leviev still runs the Newscaster. TV group and works as a producer there to air live streams. Navalny has worked with Newscaster.TV to live stream protests, such as nationwide anti-corruption demonstrations on March 26, 2017. But Leviev and Mikhailov's political activism does not necessarily taint their investigations, as they stress that "our goal is not to show the true face of wars waged by the regime [Putin's government] in Syria and Ukraine," and even seeing their work as a patriotic service to Russia, allowing greater transparency into the deadly wars that Russia has participated in since 2014. "We believe soldiers' relatives and the Russian public in general deserve accountability from military officials," Mykhailov explains. "This accountability is lacking in Russia, even when compared with the United States' less-than-ideal track record when it comes to these issues."[2]

The CIT has seen its greatest success through cooperative efforts with major news outlets, namely with Reuters (Tsvetkova 2017), SkyNews (Sparks 2016), the BBC, and Spiegel Online (2016). Leviev says that cooperating with these major media outlets is not only beneficial in increasing the quality of the investigation, but also through generating exposure throughout the Russian media landscape. "If you work together with a large publication," Leviev says, "this means that all Russian publications will write about your work – they always monitor what is being written by these major outlets, and also about what they do not cover." When a work does not come with the cooperation of a major media outlet, CIT investigations receive less attention, and have "a much weaker effect and smaller audience," as Leviev explains. However,

[1] Personal correspondence (May 14, 2017).

[2] Personal correspondence (May 14, 2017).

these cooperative efforts are not entirely without their issues. While online-based investigative outlets are able to be comparably agile in when and where they share their work, there are more stringent standards and time limits at larger news outlets, or as Leviev puts it, "The staffs at these outlets often have hard deadlines for publication, and are not able to wait around while you gather facts and continue on with investigations. They need everything *now*." While the investigations published on the CIT website or Facebook page do not receive nearly the number of eyeballs as when a report is in a Reuters article, there is more flexibility with updating an investigation with newly discovered facts.

InformNapalm

The pseudonymous Ukrainian "Roman Burko" formed the investigative collective InformNapalm following a successful period of operating his own investigative news site, BurkoNews.info. Burko, a native to the Donbas of eastern Ukraine and resident of Crimea until the Russian annexation, began writing on the Russian takeover of Crimea as it occurred, thrusting him into a new role that he had no part of in his life before war began (Nikitenko 2015). The Georgian Irakli Komakhidze co-founded the project alongside Burko in 2014 (Holub et al. 2017), with the two co-founders sharing a history of war with Russia in their home countries. Komakhidze joined Burko in the early days of investigation by helping "identify units and equipment of the Russian terrorist forces," as he described in a 2015 interview (Nikitenko 2015). Soon, InformNapalm gathered volunteers from around the world, but with a plurality of contributors from Ukraine who were concerned with Russia's involvement in the Ukrainian Conflict. There may also be some degree of familiarity or cooperation with some members of InformNapalm and the Ukrainian Security Service, judging by a report in which Roman Burko wrote that materials in an investigation were "handed over exclusively to InformNapalm from officials of the SBU [Ukrainian Security Service]" (Burko 2016).

Like the CIT, InformNapalm is almost exclusively concerned with investigating events surrounding war, including tracking military units, equipment, and sharing information on violations of the Minsk ceasefire agreement between Ukraine and separatist forces. Some of their most popular investigations have concerned Russian involvement in eastern Ukraine and alleged violations of the Minsk ceasefire agreements from

separatist forces. Materials sourced from social networks and satellite images serve as the basis for these investigations, such as "Russian cockroaches: identification of an occupier from the Russian 137th Reconnaissance Battalion in Donetsk" (Krm 2017) and "Avdiivka on fire: BM-21 GRAD MLRS position identified" (Kuznetsov 2017b). A frequent type of investigation will be to uncover the social network profiles of Russian and separatist soldiers, and cross-referencing the materials on these pages to establish which Russian military units have had active participation in the Donbas. InformNapalm also receives photographs and videos of drones launched by volunteer battalions and other organizations near the front-lines, allowing their analysts to identify types of military equipment used by the Russian-separatist forces and determine their location (Velichko 2016a).

However, in a departure from other open source research outlets, InformNapalm has embraced a number of Ukrainian hacker groups who have targeted Russian and separatist officials, providing a wealth of pilfered documents and emails concerning Russian activities in the Donbas. The most well-known of these releases were the Surkov Leaks, for which InformNapalm provided the first analysis (Velichko 2016b) after being granted exclusive access to the materials by the Ukrainian Cyber Alliance, consisting of the hacker groups Cyber Hunta, FalconsFlame, RUH8, and Trinity. These leaks received international news attention, including widespread coverage in major American news outlets (Associated Press 2016) at a time when news surrounding Ukraine was firmly on hold.

Similar to the CIT, InformNapalm has cooperated with large media companies, but its greatest successes seem to come from their cooperation with Ukrainian patriotic hacker groups. The exact extent to which InformNapalm cooperates with these groups is unclear, but it is evident that they have been able to provide detailed analysis of the information contained in the leaks, and drawn connections between data points in hacked inboxes and servers. For example, a January 2017 hack from the Ukrainian Cyber Alliance revealed documents from "the desktop computer of the reconnaissance commander" of separatist forces in Luhansk (Kuznetsov 2017a). These documents detailed the use of the Russian Orlan-10 drone in the Donbas, with direct cooperation between Russian servicemen and separatist forces. InformNapalm was able to provide analysis that further corroborated these hacks by sharing information about previous sightings of crashed Orlan-10 drones, with serial numbers and locations matching leaked documents.

Bellingcat[3]

Following years of successful blogging on the "Brown Moses Blog," UK-based Eliot Higgins launched Bellingcat on 14 July 2014 following a crowdfunding campaign that raised over £50,000. While Higgins mostly focused on the civil wars in Syria and Libya, the downing of Malaysian Airlines Flight 17 (MH17) three days after his site's launch led to the site's most well-known investigation. In the three years since the site's launch, Bellingcat has published dozens of reports investigating the open source evidence surrounding MH17, along with other topics related to the Ukrainian Conflict.

Similar to InformNapalm, Bellingcat hosts contributors from around the world with roughly equal focus on Ukraine/Russia and Syria. While the two previously discussed open source collectives focus almost exclusively on conflict-related investigations, Bellingcat uses open source research to investigate other topics, including corruption, environmental issues, and information security. With contributors spread out across the globe more diversely than with InformNapalm or the CIT, there is a greater plurality of ideologies, but the majority of published investigations take similar stances opposing the military activities of Russia and Syria, with less of a focus towards Ukraine and the West, with some exceptions.

A difference in research methodology between Bellingcat and other investigative collectives is the use of crowdsourcing the collection and analysis of information. Crowdsourcing research involves a call for readers to share the burden of data collection and verification, in hopes of increasing the quantity and quality of analysis for large amounts of information. Bellingcat's Ukraine Conflict Vehicle Tracking Project (Kivimäki 2015) produced the most success in its crowdsourcing projects, where the free platform Check (previously CheckDesk) was used to verify photographs and videos of military equipment in Russia and Ukraine. In this project, Bellingcat volunteers gathered over 350 pieces of evidence surrounding military equipment in Russia and Ukraine, with the exact times and locations established for most of these materials. Each photograph or video was collected and verified jointly between Bellingcat staff and online volunteers, establishing the veracity of the materials. For example, a video showing a tank being transported would be verified as truly in Ukraine or Russia during the time of the Ukrainian Conflict, and not

[3] Note: The Author Is a Long-Time Contributor and Staff at Bellingcat.

a recycled video from the Russia-Georgia War or the Syrian Civil War. With such a large collection of information, trends and relationships between military convoys and individual pieces of equipment could clearly emerge, such as the previously discussed example of the Lavina BMP-2.

Along with its investigations, Bellingcat has focused on providing instructional materials related to open source research to readers (Bellingcat, n.d.). Additionally, it has made efforts to establish relationships with larger media outlets through digital research training workshops. Many of the most widely-read and watched investigations into the Ukrainian Conflict include materials from open sources, making a working familiarity with the specialized skills surrounding their collection and verification essentials for journalists working in the region. In 2016 alone, Bellingcat volunteers and staff members held seventeen training workshops in seven countries, with over 200 participants representing about sixty media organizations. Among these workshops are several self-organized training sessions targeted for Russian and Ukrainian journalists, with a focus on how to successfully implement research from open source materials into their reporting.

Anonymous Sleuths and Citizen Activism

A number of anonymous digital sleuths work mostly on social media or blogging platforms, where they collect and analyse information. One of these anonymous researchers goes by the handle "Necro Mancer," or @666_mancer on Twitter, and operates a blog (donetsksite.wordpress.com). "Mancer," who lives in Donetsk, started writing in April 2014 after his home city was occupied by pro-Russian separatists. The Donetsk native's research is wide-ranging, but has two main focuses: collecting and disseminating information regarding the daily fighting throughout the Donbas from local witnesses, and collecting information on the local and foreign fighters who joined the Russia-sparked war in eastern Ukraine. Though his research methods and sources are transparent and replicable by anyone with an internet connection, he does not pretend to be neutral in how he presents his information—"Neutral truth is good," he says, "but I'd like to see Donetsk Ukrainian again, and that is my target."[4]

[4] Personal correspondence (May 14, 2017).

Askai, who also uses the pseudonyms of Askai707 and sled_vzayt, is known for his exhaustive research into Russian military equipment and units that have made their way into eastern Ukraine since 2014. The majority of his investigative work is rooted in the summer of 2014, when the largest surge of Russian resources made its way into the Donbas. The only known biographical detail about Askai is that he is a Ukrainian man, with his name, location, and background unknown to the public. Over three years after the outbreak of the Ukrainian Conflict, Askai still researches the activities of Russian forces in the hot summer of 2014, where the war turned in the favour of the self-proclaimed separatist republics over the Ukrainian Armed Forces. Unlike the three previously discussed research collectives, Askai runs a solo operation, with his personal Twitter account and a LiveJournal blog (sled_vzayt, n.d.).

Along with specialized anonymous investigators, there have also been grassroots efforts enabling locals in the Donbas to share information with Ukrainian intelligence and military officials. This trend can be seen most clearly in the use of the hashtag #StopTerror (#СтопТеррор) on Twitter in 2014, in which Ukrainians would share first-hand accounts, either through witness reports or photographic/video materials, regarding the movements of Russian/separatist military equipment and troops. Spokesperson for the Ukrainian Anti-Terrorist Operation (ATO) Press Center, Vladislav Sleeznev, praised these efforts in a July 2014 interview:

> A lot of patriotic Ukrainians are helping us by reporting on militant positions. The help of such people is invaluable. With some other people, there is also feedback—after the destruction of a firing position, these people will report to us how effective we were. If, for example, information appears in social networks, then our intelligence officials will verify this information. There have been many cases when information from local residents has been confirmed by our intelligence officials and really helped the ATO forces. (Serov 2014)

Challenges and Successes

While open source investigations into the Syrian and Libyan Civil Wars existed before the Ukrainian Conflict, the rise of open source collectives has launched the methodology to a new level of exposure and gained a degree of public trust. In particular, Bellingcat's investigation into MH17 has drawn responses from high-level officials in Ukraine and

Russia, including Russian Minister of Foreign Affairs Sergey Lavrov. Investigations from InformNapalm and the CIT have received similar attention from government officials, granting them greater clout with the public—even if the Russian officials only provide criticism of their findings.

With user-generated content and internet access expanding globally, the practices of open source research and verification will only become more important over time, both to specialized online research groups and traditional investigative journalists. The most common criticism levelled against open source investigations is the supposedly lesser value of online materials as the basis of investigations, as opposed to more traditional sources such as interviews, documents acquired from anonymous sources, and so on. However, as shown by the popularity and success of open source-based investigations into the Ukrainian Conflict, digital research can be just as trustworthy and appealing to audiences as shoe leather journalism if it is properly presented and contextualized.

References

Action Tube. (2014, August 29). *Из России идут танки с надписью "За Донбасс", "За родину"* [From Russia There Are Tanks With the Inscription "For Donbass", "For Motherland"] [Video file]. Retrieved from https://www.youtube.com/watch?v=CJm5bjM3Z5c&feature=youtu.be&t=1m27s.

Associated Press. (2016, October 26). Hackers: Emails Show Ties Between Kremlin, Ukraine rebels. *Fox News*. Retrieved from http://www.foxnews.com/world/2016/10/26/hackers-emails-show-ties-between-kremlin-ukraine-rebels.html.

Bellingcat. (n.d.). Category:Guides. *Bellingcat*. Retrieved from https://www.bellingcat.com/category/resources/how-tos/.

Bellingcat Vehicle Tracking Project. (2015, February 10). *[1140] Donbass ★ 8,000 soldados ucranianos rodeados en Debaltsevo 31.01.15* [[1140] Donbass ★ 8,000 Ukrainian soldiers surrounded on Debaltsevo 31.01.15] [Video file]. Retrieved from https://www.youtube.com/watch?v=-0DBVrKi_4k.

Bocharova, A. (2014, August 22). [Photograph]. Retrieved from https://archive.is/xSUka.

Burko, R. (2016, August 25). Doveli do psikhushki i razvalili brigadu – primery uspeshnykh informatsionnykh operatsiy [How to Drive Someone Crazy and Ruin a Brigade—Examples of Successful Information Operations]. *InformNapalm*. https://informnapalm.org/26400-doveli-do-psihushki-i-razvalili-brigadu-primery-uspeshnyh-informatsionnyh-operatsij/.

CyberBerkut. (n.d.). [Website]. Retrieved from https://cyber-berkut.org/en/olden/index8.php.

D'Agata, C. (2016, October 26). Hacking Techniques Used Against DNC, Podesta Also Targeted Journalists Covering Russia. *CBS News*. http://www.cbsnews.com/news/hacking-techniques-used-against-dnc-podesta-also-targeted-russian-journalists/.

[Digital photo]. (n.d.). Retrieved from https://017qndpynh-flywheel.netdna-ssl.com/wp-content/uploads/2015/05/lavina_10.jpg.

Holub, A., Yelyzavetta, H., Kozlikuk, S., Malko, R., & Tynchenko, Y. (2017, January). Information Napalm. *The Ukrainian Week. International Edition, 107*(1). Retrieved from http://i.tyzhden.ua/content/photoalbum/2017/01_2017/23/12/book1.pdf.

Kivimäki, V. (2015, February 3). Bellingcat Launches the Ukraine Conflict Vehicle Tracking Project. *Bellingcat*. https://www.bellingcat.com/resources/2015/02/03/ukraine-conflict-vehicle-tracking-launch/.

Kosslyn, J., & Green, Y. (2016, April 20). Montage—The Next Generation of War Reporting. *Jigsaw (Medium)*. https://medium.com/jigsaw/montage-the-next-generation-of-war-reporting-a04f4176aff.

Krm, V. (2017, March 16). Identification of an Occupier from the Russian 137th Reconnaissance Battalion in Donetsk. *InformNapalm*. Retrieved from https://informnapalm.org/en/russian-cockroaches-identification-occupier-russian-137th-reconnaissance-battalion-donetsk/.

Kuznetsov, M. (2017a, January 17). Reconnaissance Commander in the 2nd Army Corps in the Focus of the UCA. Part 1. Orlan-10 drone. *InformNapalm*. https://informnapalm.org/en/reconnaissance-commander-of-the-2nd-army-corps-in-the-focus-of-the-uca-part-1-orlan-10-drone/.

Kuznetsov, M. (2017b, February 1). Avdiivka on Fire: BM-21 GRAD MLRS Position Identified. *InformNapalm*. Retrieved from https://informnapalm.org/en/avdiivka-bm21.

Luhn, A. (2015, May 28). Vladimir Putin Declares All Russian Military Deaths State Secrets. *The Guardian*. https://www.theguardian.com/world/2015/may/28/vladimir-putin-declares-all-russian-military-deaths-state-secrets.

News-Front. (2015, February 3). *Полная зачистка Углегорска. Спецоперация ополчения. Эксклюзив. 18+* [Complete Sweep of Uglegorsk. Special Operation of the Militia. Exclusive. 18+] [Video file]. Retrieved from https://www.youtube.com/watch?v=7g05kXeULis.

Nikitenko, T. (2015, August 17). Vyzhigat napalmom. Sozdatel InformNapalm o Gosdepe, informatsionnom ugare i recepte dlya gosudarstva [Burned with Napalm. The Creator of InformNapalm on the State Department, Information Waste, and the Prescription for the State]. *Farwater*. https://farwater.net/farwater/vyzhigat-napalmom-sozdatel-informnapalm-o-gosdepe-informacionnom-ugare-i-recepte-dlya-gosudarstva/.

Roman iNapalm. (2014, December 6). *#tanks2 Енакиево 2* [#tanks2 Yenakiieve 2] [Video file]. Retrieved from https://www.youtube.com/watch?v=OQ4qbPQf0YU.

Serov, I. (2014, July 15). Mirnyye geroi Donbassa: Kak zhiteli Vostoka pomogayut armii [Peaceful Heroes of the Donbass: How the People of the East Help the Army]. *Segodnya*. http://www.segodnya.ua/regions/donetsk/mirnye-geroi-donbassa-kak-zhiteli-vostoka-pomogayut-armii-536742.html.

sled_vzayt. (n.d.). [Web log]. Retrieved from http://sled-vzayt.livejournal.com/.

Sparks, J. (2016, August 10). Revealed: Russia's 'Secret Syria Mercenaries'. *SkyNews*. Retrieved from http://news.sky.com/story/revealed-russias-secret-syria-mercenaries-10529248.

Spiegel Online. (2016, February 12). Russische Bodentruppen unterstützen Assad-Offensive. *Spiegel Online*. Retrieved from http://www.spiegel.de/politik/ausland/syrien-russische-bodentruppen-unterstuetzen-assad-offensive-a-1077054.html.

Tsvetkova, M. (2017, March 22). Russia Underplayed Losses in Recapture of Syria's Palmyra. *Reuters*. Retrieved from http://www.reuters.com/article/us-mideast-crisis-syria-russia-casualtie-idUSKBN16T0S4.

Velichko, A. (2016a, October 17). Drones Spot Another Large Concentration of Russian Weaponry in Donbas (Aerial photos, video). *InformNapalm*. Retrieved from https://informnapalm.org/en/concentration-russian-weaponry-donbas-aerial/.

Velichko, A. (2016b, October 25). SurkovLeaks: 1 GB Mail Cache Retrieved by Ukrainian Hacktivists. *InformNapalm*. Retrieved from https://informnapalm.org/en/surkovleaks/.

Vojinovic, D. (2014, September 2013). *Pro-Russian Gunmen Sit Atop Armored Personal Carrier as They Pass Through the Checkpoint Near the Town of Zhdanivka, Eastern Ukraine, Tuesday, Sept. 23, 2014* [Digital image]. Retrieved from http://www.apimages.com/metadata/Index/Ukraine/76ac286a94864a32867e324de4bcbd50/7/0.

Check for updates

CHAPTER 20

In-Depth Crisis Reporting

Natalia Antelava

The boy's name was Ziosu and he was 13 years old. I watched him as he stood in the middle of a giant pile of debris, staring at what was, only ten days ago, his family home.

It was May 2008. Ten days earlier, a massive cyclone had hit Burma's Irrawaddy Delta. We will never know for sure how many lives Cyclone Nargis claimed, but estimated hundred thousand people had died and tens of thousands went missing. In his village, Ziosu was the only survivor. When the cyclone hit, he managed to climb a tree and cling on to it for fourteen long hours as he waited for the storm to calm. From the safety of his height he watched as the dark water swallow up his mother, his father, his two brothers and his baby sister. No one came to his rescue.

The illogical cruelty of any natural disaster is difficult, perhaps even impossible to comprehend. But in Burma in 2008, a cruelty of a different sort was even more disturbing.

More than a week since the cyclone first hit, hundreds of thousands of people like Ziosu all across the Irrawady Delta were still waiting for help. They were hungry, desperate and isolated—not because of the weather,

N. Antelava (✉)
Coda Media Inc, New York, USA

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_20

but, as I discovered while covering it for the BBC, because their government was not interested in rescuing them.

Burma's military rulers had spared no effort on bringing order to the former capital Rangoon. On almost every corner of the capital there were soldiers in dark green uniforms: sweeping, fixing and cleaning. The state-controlled newspapers had been full of praise for the way the government has handled the crisis. The generals had made sure that no one was in a position to challenge their view. Army checkpoints blocked all roads to the Irrawaddy Delta. No aid workers, no foreign journalists were allowed into the country. A few who managed to get into Rangoon were deported.

I got lucky. I had an unused Myanmar visa in my passport when the cyclone hit, and I flew in pretending to be a tourist. Once in the country, I eventually met up with a BBC cameraman from Bangkok who was there under pretext of being a Thai businessman. Lying down under a back seat in a van of one very brave Burmese man, we smuggled ourselves into the depth of Irrawaddy Delta.

It was there that I met Ziosu. A lonely boy sleeping on a floor of a house of a local fisherman who gave shelter to a few survivors. When the fishermen offered to take us to the worst hit villages, Ziosu said he wanted to come along. Before dawn the following morning, we got into a small fishing boat. Shouting to each other over the rumble of an old engine, fishermen steered the boat through the heavy curtain of rain. The trail of bodies and destruction stretched along the two sides of the river. The air stunk. I counted one hundred and seven bodies before I finally gave up. They were everywhere: lining the muddy banks of the river, pushed into sludge and rubble—white, swollen figures of the countless victims of cyclone Nargis.

Across the deck from me Ziosu sat quietly, his dark eyes focused on the distance. He did not look at the bodies and paid no attention to the heavy drops of rain that fell on his face. I tried to imagine what he had just lived through and what was ahead of him: a thirteen-year-old in one of the world's most isolated countries, left without a single family member or a safety net of any sort.

As we docked, Ziosu confronted a pile of rubble that was now his home village. For a second he stared, and then, as if he had just remembered something, he jumped off and hurriedly stretched out his hand to help me off the boat.

And that is all I know about him. Not because his story ends there but because a few days later I flew out of Burma, tapes hidden in a box of tampax and I never found out what happened to him next. Because there was always a next disaster, revolution or war to cover. Because truth is, we journalists are terrible at following up.

In the news business, we always descend on places when big events happen and then leave just as suddenly, pulled out by our own circumstances rather than because the story has come to an end. The result? Gaping holes in public understanding of the crises that are shaping our world.

Had I been able to tell you what happened to Ziosu next, had I been able to tell the story of his community in the aftermath of the cyclone then I would also be in a much better position to explain to you the next big story in Burma, and the next one, too.

A few years after Cyclone Nargiz made headlines, the next time Myanmar got into news in a big way was when it suddenly opened up. It was a bizarre transformation from one of the world's worst dictatorships into a democracy. Watching it from the outside it made no sense; countries normally do not metamorphose from dictatorships overnight. But this seems to have happened with Burma, and no one could quite explain why. I got my explanation a couple of years later when I happened to be seated next to a high-ranking Burmese official at a formal dinner. I asked the man, whom I cannot name, what was behind Myanmar's incredible transformation. His answer shocked me: Cyclone Nargis.

The damage of the cyclone and the way it was handled made the generals realise that something in the system had to change, he explained. I was surprised. This simple and apparently crucial bit of information was missing entirely from a narrative that media had built around Myanmar's political transformation. What are we missing from the narrative of the following big story: the horrific persecution of tens of thousands of Rohingyas?

It is not just Burma. Did ISIS come out of nowhere? What has happened in Haiti, Syria, Libya, or Ukraine, all these places that dominate our headlines for a while and then disappear only to burst back into our living rooms with puzzling events that we struggle to understand.

The earthquake in Haiti killed hundreds of thousands in January 2010, precipitating a massive humanitarian and political crisis. Yet by October only a single foreign correspondent was based in the country to report on how the United Nations had introduced a decimating cholera

epidemic. The assassination of the US ambassador in Benghazi occurred in an on-the-ground reporting vacuum. The conflict in Eastern Ukraine erupted after most journalists had left the country following the referendum in Crimea.

We all know that context is key to understanding anything that happens in our life, be it a dispute with a neighbour, a success of a local politician, or an international conflict. Events which feed news cycles do not happen in a vacuum and crises do not end when television crews leave. In fact, the most important things often happen once journalists are gone and at times precisely because they have gone.

So why is it, then, that we journalists keep covering events as if they happen in a vacuum?

The answer is multi-layered, but it is not very complex. One reason is, of course, the decline in revenue at newsgathering organisations that has caused a well-publicised contraction in high-quality international news reporting—for many journalism outlets foreign reporting has reached a point of abdication. The reduction in foreign coverage is now a two-decade trend: Back in 1998 *The American Journalism Review* identified 18 newspapers and two chains which had shuttered every one of their overseas bureaus during the previous twelve years (Arnett 1998). Since then, the retreat from overseas coverage has only worsened even if media companies like to pretend otherwise. By 2012 *The Los Angeles Times* claimed 10 overseas "bureaux," but eight of those bureaus consisted only of a single reporter (Martin 2012).

Another, often sighted reason behind journalism's lack of follow up is the modern-day panacea of short attention spans. There are, of course, plenty of fantastically dedicated reporters covering all sorts of issues, and most of them know how hard it is to get editors' attention once the story they are covering has lost its place in the global running order.

And yet it is not the attention span or even lack of resources that really prevents us journalists from following up. The reason why we are so bad at giving our audiences a meaningful follow up is because journalism, as we know it, is simply not designed for staying on a story.

Traditionally, journalism has always been housed on disposable platforms: Think of newspapers that go into a dusty pile or television or radio pieces that vanish after being aired. All traditional journalism formats that we know are disposable, which, for decades, has made it nearly impossible for reporters to create a meaningful follow up.

Take the story of Ziosu as an example. If I had spent ten months reporting on Ziosu's story for a conventional newspaper, I would have needed to make sure that my readers had access to every preceding story that I had written. But because they would not, as a reporter I would have needed to tell the story from scratch each time, with basic facts and background repeated each time and only limited space left for an actual update. Essentially, traditional formats force journalists into dumbing down stories as we can never assume that our readers or viewers have the background needed to understand the update.

Structure is at the heart of any good storytelling; journalists know the value of constructing narratives that keep people reading, listening and viewing. However, the digital age has rendered our ability to tell a good individual story insufficient. The traditional news story is a self-contained item with a headline, picture and text, which is repeated every time there is a development just in case the reader missed previous edition. When published online, this format creates an overload of information, with millions of fragmented versions of the same story and an absence of cohesive narrative.

Internet, however, has fundamentally changed the disposable nature of journalism platforms. It gave us an opportunity to create digital scrapbooks, where each individual story can be told in wider context, allowing us to follow characters, events and trends over a period of time. Digital formats now allow us to show how stories relate to each other; they empower journalists to place intimate, engaging, on-the-ground storytelling into a much wider context of crises and events.

It is a revolutionary change, but one that the industry has been very slow to embrace. When newspapers and broadcasters created their new homes for themselves online, they essentially copy pasted the format that they were familiar with. Internet became a digital pit for endless updates, a virtual equivalent of a pile of newspapers that we used to see in our parents' hallway: full of invaluable information, lessons of the past, and clues to the future, all hidden away in a plain site.

In fact, internet amplified the inherent superficiality of a news cycle by introducing a rolling deadline. For most journalists, there was no longer an evening broadcast or a morning paper to aim for; instead, there are constant updates to be filed anytime day or night. The constant pressure to file not only increases a reporter's workload, it also interferes with a journalist's ability to report. The more time we spend filing, the less time we have left to make phone calls, to travel, to talk to people directly

affected by a crisis. The more we file, the more locked in we become in an echo chamber, repeating and regurgitating what has already been said.

I remember landing in Sanaa, Yemen, in 2013 as the country was being shaken by massive and increasingly violent anti-government protests. My phone rang just as I cleared security. An editor in London wanted me to send a dispatch as soon as possible. "But I just arrived," I tried to argue. It did not matter—the beast had to be fed and the fact that the beast was now digital made its appetite insatiable.

Six weeks later, in the departures lounge in the same airport in Sanaa, I wondered whether there was another, better way of covering events in digital age. Walking towards my gate, I thought of Yemen's uncertain future, the people I was leaving behind and their stories that would now go untold. I had another assignment waiting, and it is not that my editors did not care about Yemen; but with resources scarce and other stories fighting for their attention, the decision had been made not to send in another reporter. I knew that without a journalist pushing the story from the ground, Yemen would disappear from the editorial agenda, just like Burma or so many other stories I had covered have slipped down newsroom's running orders in the past. I thought of Ziosu and wondered what had happened to him.

After my trip to Yemen I reached out to several colleagues who I knew shared my frustration with the way media covers big crises. Together with the American journalist and fellow veteran foreign correspondent Ilan Greenberg we embarked on a mission to figure out a model that would allow journalists to take stories out of an erratic news cycle and cover crises in a way that creates a meaningful, cohesive narrative.

Coda, the journalism start-up that we co-founded, was born out of a three-year long conversation that we initiated with a group of young and veteran reporters, designers, technologists and editors with whom we skyped, met and corresponded between deadlines and assignments.

Coda is a musical term—and in music it is a passage, usually at the end of a piece, that helps define the whole work. In journalism, we decided, Coda Story would be a stand-alone voice that helps define a whole crisis.

We set our goals: to investigate crises that shape our world, to challenge conventional narratives, to make important, complex stories accessible to general public, and to promote in-depth journalism through partnerships, collaborations and mentorship schemes. As a non-profit,

Coda would also support freedom of the press, without which consequential journalism cannot exist.

We quickly realised that in order for us to achieve any of the above, to stay on the story, and to keep our audience on the story with us, we could not simply continue to produce coverage when others had left. We had to do something more fundamental: to move away from news as a constant stream of incrementally-updated articles and to stop using the internet as a bottomless pit for updates. We had to re-imagine journalistic storytelling in digital age by moving away from the mentality of a disposable platform.

A successful Coda, we thought, could not be a news dashboard, aggregation site, user-generated platform, or hyper-specialist watering hole. It would need to produce high-quality, original narrative storytelling that attracts a wide readership via traditional text, features and video, as well as distinct, innovative and inventive digital formats. Coda's success would lie in its ability to yield a complex, nuanced, and contextual understanding of specific storylines fuelling contemporary global issues.

The operational formula we created to achieve this was simple: Coda, a non-profit newsroom, would deploy like everyone else after a crisis, but go in with a long-lens view and commit to a story for an extended period. The core editorial team would then define trends or "currents" that run through the crisis and works with a network of international and local journalists to report on different aspects of the issue. Each individual piece of content, regardless of its format, would be placed in one of the trends or "currents" allowing our audience to follow a narrative line and see the wider context. Currents, in other words, would be the heart of Coda's storytelling; a fusion of editorial and design enabling connection and context to work as hard as the individual pieces of content.

To test Coda's editorial concept, in 2016, we ran a successful crowdfunding campaign, matched it with a small foundation grant and launched an MVP, a minimal viable product.

For our pilot project, we picked the subject of gay rights in the former Soviet Union. Media widely reported on gay rights abuses around the time of the Sochi Olympics in 2014 but then moved on, largely abandoning the subject. However, by 2016, from the Baltics to Central Asia the letters LGBT were no longer just an expression of identity. They had instead turned into a rallying cry, either as a fundamental benchmark of tolerance or as a hate phrase defining a threat to tradition.

Our aim with the pilot was not just to follow up on individual stories of homophobic abuse and violence but place them into wider context of geopolitical developments. We did not just want to show that gay men and women were victims, we wanted to understand why they were victims, and what the reasons were the Kremlin had chosen to wage its war on gays. Through dispatches, video web-series and photography we told the story not only of individuals affected by this crisis but of how the crisis of gay rights pit the Kremlin against the West, both reflecting and feeding the wider geopolitical story, whether in Syria, Ukraine or among the right-wing politicians in the United States.

The pilot edition explored the crisis by following five currents, or trends, that ran through it: In the *Rights Abuses* current we tracked the human cost of cultural and state-sanctioned homophobia and reported on individual cases of violence and abuse. In the *Kremlin Influence* current we examined Moscow's role in turning sexual orientation into a matter of political ideology not only in Russia but across the former Soviet Union, while in the *Orthodox Church* current we told stories of the central role the Church has played in framing this new "traditionalist" ideology. The stories in the *East-West* current tracked the growing divide between two schools of thought as Moscow's definition of values and tradition clashed with the Western emphasis on individual rights and, finally, stories in the *Information Wars* current followed coverage patterns and decoded agendas behind them. The content was disseminated through Coda's own social media and published by a number of editorial partners, including *The Guardian*, World Policy Journal and Reveal, a podcast from the Centre for Investigative Reporting.

There were three main key takeaways from our test pilot edition. The first was that such editorial framing of a crisis and thematic rather than chronological curation of the content did not just allow us to follow trends and individual characters as their stories developed, but crucially it generated stories that had an infinitely longer shelf life than those reported in reaction to news events.

The second lesson was that when selected carefully, currents could greatly aid the process of journalistic discovery. For example, a current that unexpectedly emerged as most popular and hence most populated in our pilot edition was the one tracking Information Wars. As the editorial team dug into the crisis, commissioning freelancers from across the region and following the news, we realised that the role of

disinformation and propaganda against gays was not just key to understanding the issue, but it had also been severely underreported.

Disinformation soon ballooned into its own Coda edition, which we in turn broke up into individual themes. It led us to covering another crisis connected to Disinformation: that of migration in Europe. We continue to develop editorial offshoots of our initial subjects, linking crises and big themes, creating a cohesive narrative of the modern world where each individual story works as a building block of a wider, overarching narrative.

The third lesson of our pilot edition was just as consequential. While design and editorial architecture has allowed us to create a system in which journalists can stay on the story, there is only one thing, we learned, that will keep our audiences on the story with us: outstanding storytelling.

Innovative storytelling is one of the buzzwords of the media industry, but we have learned to be cautious about using the term because at Coda, we, in fact, aspire to what some would call "old fashioned" journalism, in which reporters are free from filing constant updates and instead are encouraged to dig deep, look beyond social media to find voices that are not being heard and help audiences connect the dots often obscured by the drama of breaking news. The new formats such as 360° video or animation, we learned, can be empowered when coupled with traditional values of journalism: accuracy, fairness and powerful, character-driven narrative storytelling.

In 2013, Pew Research published a study that offers insight into a journalism gap—the unaddressed desire among news consumers for quality storytelling predicated on deep reporting in the field (Mitchell et al. 2013). The study found that longform journalism stories which contain deeper reporting or analysis and tend to be more than 1000 words long appeared infrequently in the start-up media that Pew evaluated. In a two-week period, about two-thirds of the outlets did not produce a single longform piece and most of the remainder only produced between one and five longform pieces.

News consumers are not monolithic, and many are indeed the peripatetic readers of nugget-sized news flashes and RSS headlines. But there are also legions of readers interested in better, slower, more detailed storytelling that offers meaningful insight into the issues and events that go into invisible remission in the major media only to suddenly re-emerge

like multiplying cells, an unexamined tumour affecting their economies, their national security, their place in the world.

There is an emerging recognition of a strong appetite for journalism predicated on sustained, longform storytelling. New feature departments at legacy media organisations like *The Guardian* and initiatives from newcomers like *Buzzfeed*, *Politico,* and *Medium* reflect how investment in the longform genre is on the uptick.[1] Their efforts at deeper excavation, however, are spread across topics, regions, everything and everywhere; their individual stories illuminate a subject for a moment. There are also other news start-ups, like Coda, which chose one topic to focus on.

In fact, the 2013 Pew Research study also found most of the 73 recently founded journalism non-profits it surveyed have niche orientations—they focus on one topic or type of reporting. Only 26% of the 172 non-profits cover general interest news (Mitchell et al. 2013). The bulk of them focus on narrower topics to fill one gap or another. Other niche areas include government, health, education, and the environment, with foreign affairs occupying a focus of only 13%. While "nichification" seeks to fill a hole in newsgathering, media start-ups focusing on international stories remain few.

The growth of single-issue media start-ups and the efforts of many media organisations to build better topics pages as well as explore deeper forms of storytelling are indicators of the demand for deep, specialist information. Coda's single-subject agenda can uniquely lash the power of longform journalism to a topic over time, creating a unique digital record of an issue or a major event, and eventually scaling up by deploying on several crises at the same time, creating a niche audience with each deployment.

Will it work? There will be plenty of sceptics who would argue that in the age of short attention spans, people would never want to stay on a story of a faraway place. We believe it depends how it is told. Would you not want to hear what happened next to Ziosu, a boy who had lost his entire family and yet thought of jumping off a boat first in order to help a stranger?

[1]For an analysis on key drivers of the resurgence in longform journalism, see: http://pando.com/2013/08/12/epic-launches-politico-goes-deeper-why-longform-is-flavor-of-the-month/.

References

Arnett, P. (1998, November). State of the American Newspaper. Goodbye, World. *American Journalism Review*. http://ajrarchive.org/Article.asp?id=3288.

Martin, J. D. (2012, April 23). Loneliness at the Foreign 'Bureau'. *Columbia Journalism Review*. http://archives.cjr.org/behind_the_news/loneliness_at_the_foreign_bureau.php.

Mitchell, A., Jurkowitz, M., Holcomb, J., Enda, J., & Anderson, M. (2013, June 10). The Landscape. *Pew Research Center*. http://www.journalism.org/2013/06/10/landscape/.

Check for updates

CHAPTER 21

eJihad: Behind the Use of Social Media by ISIS

Asiem El Difraoui with Oliver Hahn

Introduction

Today's so-called eJihad is a worldwide phenomenon and the cross-media propaganda toolbox of the global Jihadist movement. Without the audio-visual propaganda to unite sympathisers by creating a shared identity, the global Jihadist movement might no longer exist (El Difraoui 2015, p. 127). Jihadism can generally be understood as the most radical interpretation and distortion of Islam and its concepts as well as the most extremist current and movement of Islamism within Sunni Islam. In classical Sunni thought, the concept of Jihad has many different meanings—such as the struggle within oneself to become a better human being. Jihadists, however, reduce the concept to armed struggle and violently attempt to impose their political and societal worldview on others. Fighting is not only considered the most important instrument for solving the problems of the Muslim world, but also for personal salvation including the supposed martyrdom through

A. El Difraoui (✉)
Institut für Medien- und Kommunikationspolitik, Cologne, Germany

O. Hahn
Centre for Media and Communication, University of Passau, Passau, Germany

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_21

suicide attacks. It is important to stress that—within Sunni Islam—three different forms and groups of Islamism: political Islam(ism), Salafism (this current is strongly influenced by the long-time Saudi Arabian state doctrine of Wahhabism) and finally the Jihadism, have to be differentiated. All forms of Islamism also have a very strong transnational component. The diversity of these different groups is reflected in the extreme variety of the Islamist media landscape. It is equally important to stress that a large part of Muslims all over the world are not Islamists (El Difraoui 2015, pp. 117–118).

The emergence of the 'Arab Spring' in late 2010 and early 2011 allowed numerous actors, groups and parties of especially political Islam(ism) and Salafism to operate media for the first time or more and more officially (El Difraoui and Hahn 2013; El Difraoui 2015, p. 122). The phenomenon of eJihad is part of the rapid development of Islamist media. Also, eJihad remains a key recruitment tool for attracting young people to pseudo-religious terror on a global scale (El Difraoui 2015, p. 128).

This chapter focuses on how eJihad has developed since the late 1970s and specifically traces how it has expanded in reach and diversity since the rise of social media. The history and proliferation of Jihadist propaganda can be roughly divided into four phases: its prehistory in the Afghan war against the Soviet Union (1979–1989), the spread of propaganda efforts to the Western world and the appearance of the first Jihadi websites (1990–2001), the globalisation of the Jihadist internet (2002–2006) and finally the Jihadists' use of social media (2006 to the present). The fourth and current phase of eJihad is characterised by its use of cross-media elements and has grown into a flood of Jihadist propaganda since 2014 through the activities of the so-called Islamic State (ISIS or ISIL or Daesh). The propaganda campaign has peaked in 2015 and 2016, and it has somehow receded in 2018 after military defeats of ISIS, but it could anytime and anywhere in the world gain again in strength due to the large number of global sympathisers and affiliated organisations. The propaganda generally merges Jihadist symbolism with the language and imagery of global youth culture and the extreme brutality of some video games to create a Jihadist subculture that is increasingly attractive to Europe's youth (El Difraoui 2015, p. 123).

Phase One: Prehistory in the Afghan War against the Soviet Union (1979–1989)

During the first phase, the war in Afghanistan, Jihadist propaganda was disseminated primarily through traditional print media, audiocassettes and the first (and back then expensive) VHS videotapes. Abdallah Azzam (who is often viewed as the ideological founder of Jihadism) published a magazine called *Jihad* to report on the 'glorious' deeds of the Afghan Mujahedin and the brutal acts of the 'Russians' to successfully recruit followers and solicit donations. The first videos about 'martyrs' also began to appear around this time. The fact that Jihadist cameramen who lost their lives were considered 'martyrs' just as the fighters were, shows that propaganda was already then thought to play a decisive role. Influential theoreticians such as Abu Musab Al-Suri later criticised these propaganda efforts as insufficient because they were only reaching Jihadi circles and not the 'Umma', the Muslim community in general (El Difraoui 2015, pp. 123–124).

Phase Two: The Spread of Propaganda Efforts to the Western World and the Appearance of the First Jihadi Websites (1990–2001)

In the second phase, when propaganda was spreading globally, the first elements of a Jihadist communications strategy began to emerge. The propaganda became more professional and took hold in Europe. These developments began in the early 1990s in what is referred to as "Londonistan" (Thomas 2003). At the time, the United Kingdom's renown freedom of press and opinion provided Islamists of all kind, including former Afghan fighters who had sought refuge in Britain, a unique platform for their propaganda—including calls for Jihad. The Bosnian War, which began in 1992, served as a catalyst for their efforts. The suffering of the Bosnian Muslims generated a huge wave of sympathy among the Muslim diaspora in Europe, which the propagandists used to their advantage. Influential texts in terms of Jihadi propaganda strategy appeared. Abu Musab Al-Suri called for a diversified propaganda tailored to the various target audiences. Using the motto 'system, not an organisation', he called for a decentralised form of information warfare in which independent organisational groups use leading technology to

promote a Jihadist ideology and serve as a blueprint for future propaganda. In 1996, shortly after the Bosnian War ended, the first Jihadist website, Azzam(dot)com, appeared. This website was developed by a 24-year old computer science student at London's Imperial College by using the college's server and served as a prototype (El Difraoui 2013a; 2015, p. 124).

Phase Three: The Globalisation of the Jihadist Internet (2002–2006)

In the third phase commencing after September 11, 2001, the subsequent intervention in Afghanistan and especially the start of the Iraq War in 2003 provided the Jihadists with numerous new sympathisers, and eJihad became a global phenomenon. The unpopularity of the American invasion of Iraq among Muslims allowed the Jihad propagandists to attract thousands of new recruits, and to achieve important successes in their media war. Four key developments have spread the Jihad propaganda globally. First, a considerable amount of the propaganda material no longer stems directly from the strict hierarchy of the Jihadist organisations and their media committees, but is rather disseminated directly to interested persons by decentralised so-called independent media companies. Secondly, traditional websites (in which interaction between the producers and consumers of information was rarely possible) have been replaced by dynamic, interactive online forums. As a third factor, technological advances (such as inexpensive digital cameras and video-editing software for laptops) led to an exponential growth in video production. The fourth factor was the emergence of propaganda products directly targeted at the youth influenced by global pop culture in Muslim countries and Europe (El Difraoui 2012). In online forums, Jihadist sympathisers posted materials (and especially existing materials from Jihadist organisations) for worldwide dissemination. Among these media companies, one in particular played a key role: the 'Global Islamic Media Front (GIMF)'. Generally, the Jihad's communication platforms were characterised by, among other things, a very rapid expansion throughout other languages. Nearly every European language, from Swedish to Spanish, can be found in the Jihad's online presence. This has enabled the Jihad's media staff to address local issues and alleged discrimination against Muslims in the local language and context. Examples can

be found in texts addressing the Mohammed caricatures in Denmark or the banning of headscarves in France. Through the use of online forums, the Jihadist internet has metamorphosed into an essentially non-hierarchical yet widely interconnected propaganda network. This has enabled a much broader dissemination of information and laid the foundation for the production of the videos that have played such a key role in the emergence of a Jihadist subculture or anti-culture in Western nations. Jihadist symbols and content have been adapted to the global youth pop culture in an attempt to appeal to the youth in the West. A so-called 'Al-Qaeda Rap' began to appear as early as 2005 in the music video "Dirty Infidels". These productions were the starting point of the fourth phase as well as the on-going propaganda campaign launched by ISIS in 2013 (El Difraoui 2013b; 2015, pp. 124–125).

Phase Four: The Jihadists' Use of Social Media (2006 to the Present)

The fourth phase of eJihad, introduced by the emergence of the Web 2.0, continues until today. It has been characterised by three key developments. First, the Jihad propagandists began to expand their activities into social networks and the ubiquitous smartphones. This allowed a form of Jihadi youth subculture and anti-culture ("Jihadi Cool"), to expand into the Western world. Secondly, the Jihadists tried to hijack the Islamic world's mainstream forums for their purposes. Thirdly, the Jihadists strengthened their activities in the 'deep dark net', the hidden internet. Social networks such as Facebook, together with video portals such as YouTube and message services such as Twitter, have evolved into important propaganda tools and a stronger interconnection of the group's cross-media activities emerged: a Twitter tweet mentions a new YouTube video or a Facebook page, and vice versa. The content of social media platforms is furthermore often re-used in traditional media or propaganda tools such as Jihadist print publications or flyers. The rapid, direct and global communication via various channels and the emotion-laden multimedia content consisting of personal narratives, text, music and videos are designed to give users the feeling of belonging to an international Jihadi community or Jihadist subculture which has a status equal to other communities and cultures in the digital media (El Difraoui 2015, pp. 125–126). The Jihadist have since 2016

increasingly shifted to encrypted social networks and messaging services like Telegram and WhatsApp to avoid detection and censorship.

A German-Ghanaian former rapper Deso Dogg (born Denis Cuspert) who began calling himself Abu Maleeq after his conversion to Islam, is considered to be one of the most prominent propagandists of the Jihadist subculture. While he is assumed dead, he still has a strong presence on YouTube and Facebook through his videos. Digital natives who rap (such as Deso Dogg) and thousands of other radicalised Europeans use the codes, gestures and aesthetics of European and US subcultures to promote the eJihad's propaganda and symbolism to influence a new generation of powerful propagandists and spread global terror. For example, a universally recognisable symbol in the Jihadist community is the raised index finger, whether in a selfie showing the subject posing in front of a tank or in their own living room. The gesture originated in the Web 2.0 repertoire of members of this community and can be seen on hundreds of photos and videos from the so-called Islamic State. The gesture is intended to remind Salafists and Jihadists of the important concept of 'tawhid', the unity of God and the 'Umma'. In reality it is a new symbol, the Jihadist version of the thumbs-up 'Like' on Facebook and symbol of group identification not unlike certain gestures of street gangs (El Difraoui 2015, p. 126).

Whole legions of 'lolcats' are also part of Jihad 2.0's new repertoire. The cat was purportedly one of the Prophet Mohammed's favourite animals. According to historical sources, anyone mistreating a cat is condemned to the torment of purgatory. In Islamic history, however, the symbolism of the cat is completely new. The images of cat-caressing fighters flooding the social networks seem intended to portray the fighters as 'the nice guys next door' and present them in a sympathetic light. The Jihadists are thereby part of the global trend flooding the web with cat selfies and videos (El Difraoui 2015, p. 126).

ISIS's 'virtual Jihadistan' is simultaneously a world filled with unbridled horror and new levels of barbaric violence. The 'death sentence' and brutal beheading of the American Nicholas Berg in 2004 by the mythical founder of the ISIS movement and 'Butcher of Bagdad', Abu Musab Al-Zarqawi, was still filmed using a stationary camera similar to Iraqi TV broadcasts of communiqués during the reign of Saddam Hussein. More recently in a general trend and professionalised way, more and more Hollywood-style audio-visual codes are to be found in ISIS's Jihadist propaganda. The perverse video of the murder of US journalists, Steven

Sotloff and James Foley, in 2014, dressed in orange prison garb, is reminiscent of the final scenes of Brad Pitt in the movie thriller "Seven". The threatening speech by the British ex-rapper Jihadi John after the beheading was filmed with a handheld steady-cam using imagery similar to that found in music videos (El Difraoui 2015, p. 126).

The decision taken in this context by some Western media outlets to no longer report and by doing so, to no longer further distribute ISIS's Jihadist propaganda, but instead to simply show the black-and-white flag and logo of Daesh, has become a controversy. In a way, showing the logo endorses the theft of an Islamic symbol by Daesh: in fact, the logo itself resembles the seal of the prophet.

Conclusions and Perspectives

Jihadists have managed to reinterpret and hijack Islamic concepts and symbols to create their own imagery and ultimately a Jihadi subculture. The elimination of language barriers has greatly expanded the propaganda's reach. Today's youth now finds it easier than ever to identify with the Jihadists as role models, making it even simpler for them to overcome their inhibitions and join a Jihadist organisation. All Jihadist propaganda products reduce complex theological content to the level of rhetorical dichotomies consisting of 'us' (the 'true' Muslims) vs. 'them' (everyone else) in which there can be no compromises. The goals of Jihadist propaganda are complex and multifaceted. At the base level, the goal is to attract a maximum amount of media attention and thereby recruit young Muslims and to obtain financial, material and physical support to the Jihadist cause. At the second level, the Jihadist movement tries to create a closed worldview and to convince the Muslim community that it (the Jihadist movement) alone represents the true faith. It thereby strives to portray itself as the sole authority on key Islamic concepts and symbols, an attempt which has been successful to a certain extent. On the third level, the Jihadists attempt to manipulate these concepts and symbols to establish their own mythology and doctrine of salvation. This eschatology encompasses a cosmology in which the Jihadists are the only true believers and in which only they ascend to paradise. The leaders of these Jihadist groups are portrayed as the only ones authorised to preach divine revelation and become self-proclaimed prophets through the promise of salvation accorded to simple combatants through their martyrdom. Representatives from all Islamist

currents—political Islam(ism), Salafism and finally the Jihadism—will continue to try to expand their media presence and use all technologies available to them to their own advantage. In light of these factors, it becomes clear that the Jihadist media is an inflammatory propaganda tool. Global education and prevention campaigns as well as counter-narratives have been launched already with varied success (El Difraoui 2015, pp. 127–128).

Western and European countries in particular but also several Arab nations still face the challenge to offer their own 'grand-narratives' to counter the Jihadists' one via digital media. Innovations in digital investigative journalism on an international level can provide fact-checked information and data that make propaganda and information warfare more transparent to audiences, and that, at the same time, challenge the Jihadi narrative.

References

El Difraoui, A. (2012). *jihad.de. Jihadistische Online-Propaganda: Empfehlungen für Gegenmaßnahmen in Deutschland*. SWP-Studien S 05, February.

El Difraoui, A. (2013a). *Al Quaida par l'image. La prophétie du martyre*. Paris: Presses Universitaires de France.

El Difraoui, A. (2013b). Propaganda und Märtyrertum: Drei Jahrzehnte Videodschihad. *Blätter für deutsche und internationale Politik, 6,* 43–51.

El Difraoui, A. (2015). Islamistische Medien: Vom Wahhabismus über die Muslimbrüder zum Cyber-Dschihad. In C. Richter & A. El Difraoui (Eds.), *Arabische Medien* (pp. 117–128). Constance: UVK.

El Difraoui, A., & Hahn, O. (2013). Der arabische Schwarm im Netz des Medienwandels. Ein Salon-Streitgespräch zur Protestkommunikation ziviler Bewegungen in mehr als 140 Zeichen. In E. Bettermann & R. Grätz (Eds.), *Digitale Herausforderung. Internationale Beziehungen in Zeiten von Web 2.0* (pp. 111–118). Göttingen: Steidl.

Thomas, D. (2003). *Le Londonistan. La voix du djihad*. Paris: Editions Michalon.

Check for updates

CHAPTER 22

Truth Corrupted: The Role of Fact-Based Journalism in a Post-Truth Society

Florian Stalph

The Pulitzer Prize winning fact-checking website PolitiFact found that out of 489 statements made by the 45th president of the United States during his presidential campaign and while he has been in office, 34% are false and another 15% are "pants on fire" false (as of 08/01/2018), meaning that the accounts are not only wrong but also add untenable assertions.[1] While this might not come as a surprise and false claims have been swiftly corrected by legacy media outlets over the last years, we must acknowledge that almost half of all statements by Donald Trump are mostly true, half true, or mostly false. Having a factual core, therefore portraying at least some 'truth' or providing an accurate piece of information, calling out these statements is more difficult than such being completely made up. On April 7, 2017 the US president claimed a 64% reduction on illegal immigration on the southern border of the

[1] For a current overview see http://www.politifact.com/personalities/donald-trump/.

F. Stalph (✉)
Centre for Media and Communication, University of Passau, Passau, Germany

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_22

US in the previous month. In the months before, he cited other statistics, highlighting previous decreases on a sequential monthly basis. This time, however, he did not choose to quote month over month measurements but year over year measurements—without declaring it. Monthly numbers would have suggested a 35% decrease and while these numbers were apparently too low they would have at least shown measurements comparable to those announced through previous statements (Valverde 2017). Not only did the president deliberately change the measurement category without indication—he chose statistics that fit his narrative while completely breaking with his previously established use of statistical comparisons. As a result, this inconsistent use of data reveals a systematic use of numbers that favour a certain agenda. Furthermore, through focusing on the southern border, the fact that illegal immigration via the northern border has been on a rise was neglected. A prime example that "politicians have leaned heavily on statistics to buttress their authority. Often, they lean too heavily, stretching evidence too far, interpreting data too loosely, to serve their cause" (Davies 2017).

Just as PolitiFact tries to evaluate statements made by American political actors, the collaborative effort CrossCheck has been debunking dubious (social) media claims by bundling the power of 37 newsrooms in order to call out political meddling in the French presidential election by correcting disinformation and misinformation. During that time they could identify 10 misleading and 19 manufactured stories and dozens of other misreported, manipulated or misattributed stories (CrossCheck 2017). Most of these indicted reports stem from news websites or blogs and are shared via Twitter or disseminated through Facebook sites, and most of these sources can clearly be affiliated with certain political movements as they clearly propagate disinformation aimed at harming political enemies through establishing toxic narratives. These platforms are only two examples for fact-checking organisations that set out to defend objectivity and transparency. Voters and societies in general are in dire need of such institutions around the globe. The US and French election, Brexit, the Russian insurgency on Ukrainian soil and preceding information war during the Euromaidan, the Four Day War between Armenia and Azerbaijan—recent years are stacked with episodes of disinformation and misleading reporting putting a spotlight on media organisations and social media being subject to manipulation and exploitation as means of propaganda.

The Duke Reporter's Lab has been observing fact-checking projects around the world since 2014. Whereas they counted 44 active sites in 2014 (Adair 2014), they report 126 active organisations in 49 countries three years later (Stencel 2017). More interestingly, fact-checking appears to become a global movement: Africa, Asia, Australia, Europe, North America and South America all have agencies "driven by concerns about a global epidemic of misinformation, viral hoaxes and official lying" (Stencel 2017). The Duke Reporter's Lab further determines that more than half of these projects have ties to media organisations while others are affiliated to journalism research institutes, universities or non-governmental organisations. When analysing the roots of the global fact-checking movement, Graves (2018, p. 618) proposes a model in order to map the international landscape of fact-checkers by locating them in a triad composed of journalism, academia, and politics and civil society according to values the organisations represent and affiliations they have. He concludes that fact-checking is a highly diversified system being a "forceful example of non-news organisations adopting an emerging journalistic genre in order to pursue a civic agenda" (Graves 2018, p. 624). Despite persistent differences between projects—differing organisational and methodological approaches or varying values and institutional affiliations—transnational collaboration flourishes (Graves 2018, p. 627).

The rise of fact-checking movements around the world is symptomatic of shifting paradigms in post-truth societies: Information dissemination and information consumption have fundamentally changed.

Disruptive technologies provide new means of communication that circumvent traditional communication models and turn established media logics upside down. By providing new channels and platforms that allow spreading information without verification and without institutionalised ministration, such technologies evade correctives that advocate truthful and accurate information flow, predominantly journalistic institutions such as TV and radio broadcasters, newspapers, news magazines and internet-based news media. As information now flows more freely than ever before, new players appear, filling gaps created by innovative disruptions. Disrupted information dissemination offers a wide array of options to political and societal influences from all sides of the political and social spectrum, having manifold implications for the circulation of news and messages. For instance, these new logics allow breaking through institutionalised barriers set up by autocratic regimes in order to impede public communication and to diminish alternative voices; at the

same time, political actors, interest groups or subversive powers can use certain channels and platforms to undermine the authority of political elites and leaders as well as renowned media organisations. In the end, gate keeping mechanisms seem to be overridden.

Kurt Lewin was the first to coin the term "gate keeper". Focusing on group dynamics, the psychologist analysed food habits of Americans after the Second World War. Lewin (1947) was looking into the question of who is responsible for dishes on American families' tables. He found that food components pass through several sections which are affected by certain powers. Certain areas in each channel constitute gates which serve as partitions of forces and can completely alter them, even reverse them. While he found that housewives—who Lewin identified as the most likely gate keepers of buying food—determine what food is served, what ingredients are bought, how much money is spent on them and ultimately what the members of a group, that is a family, eat, he also claimed that this concept "holds not only for food channels but also for the travelling of a news item through certain communication channels in a group" (p. 145). White (1950) put this to test and transferred Lewin's theory into a journalistic context. He observed a wire editor as the gate keeper who determines which news items lastly appear on the pages of a newspaper. White stated that the news selection is highly dependent on the editor aka the gate keeper, his own experiences and mind-set, his own evaluation of the newsworthiness of a story, and his estimation of the paper's audience. At last, White concludes that the process of gate keeping is highly subjective:

> It is a well known fact in individual psychology that people tend to perceive as true only those happenings which fit into their own beliefs concerning what is likely to happen …. in his position as 'gate keeper' the newspaper editor sees to it (even though he may never be consciously aware of it) that the community shall hear as a fact only those events which the newsman, as the representative of his culture, believes to be true. (White 1950, p. 390)

White describes a phenomenon comprehensively discussed by the social psychology community. The cognitive dissonance theory was put forward by Festinger (1957). When he analysed the interplay between human actions and underlying beliefs and attitudes, he established that people adapt their set of beliefs and ultimately act

accordingly when a situation arises that causes dissonance with pre-existing attitudes. As a premise, individuals are perceptive to inconsistencies that are caused by a conflict between current beliefs and new, contradictory information. Now, in order to act in harmony with one's attitudes, eliminating a discomforting situation, the individual seeks to resolve or reduce the dissonance by choosing to change behavioural or environmental elements by avoiding certain situations and information that are likely to cause cognitive dissonance, thus, selecting information that matches one's beliefs or behavioural elements. This concept—selective exposure—describes the subjectivity of a gate keeper's actions. Building on cognitive dissonance and substantiating selective exposure, Klapper (1960) determines that media do not change recipients' behaviours or beliefs rather than strengthening them. Within a series of experiments, Brock and Balloun (1967, pp. 424–425) found that there are "relationships between behavioral commitment and behavioral self-exposure to pertinent propaganda". Participants had to press a button to remove a static from spoken messages in order to improve reception. Their studies show that smokers were found to be more interested in messages refuting a connection between smoking and lung cancer while being disinterested in messages that uphold cancer is caused by smoking. Moreover, the amount of button presses positively correlated with the amount of cigarettes smoked. More recent studies have shown similar effects. A pilot study by the YouGov-Cambridge Programme explored how British citizens are affected by conspiracy theories. One noteworthy result is that conservative UKIP voters are more likely to believe that Brussels and the EU are plotting to take power away from the United Kingdom. Liberal Labour voters, however, are less likely to believe that climate change is a hoax (Rogers 2015). Another poll shows that almost half of Trump supporters completely distrust economic data provided by the government whereas only 5% of Clinton supporters do so (Ryssdal 2016). These examples show that beliefs of citizens oftentimes align with ideologies promoted by affiliated political parties.

The early psychological experiments have shown that information consumption is selective: Supporting information is more likely to be sought out, while information that is not congruent with one's views is avoided. As a result, the shaping of public opinion appears to be subject to the combination of (a) disruptive technologies that circumvent established logics and redefine information flows by debilitating force fields around gates, (b) recipients' behavioural patterns explaining deliberate

devotion to self-affirming information and negligence of deviating information, and (c) political powers that seek to benefit from new communication patterns and information flows, promoted by political campaigns that systematically utilise this new environment—an interplay of these factors, rather than regarding them as well-defined effects.

Former US President Barack Obama spearheaded the use of social media tools to directly address citizens, promote himself or raise donations for his election campaign. His internet strategy turned out to be highly successful—according to Arianna Huffington: "Were it not for the Internet, Barack Obama would not be president. Were it not for the Internet, Barack Obama would not have been the nominee" (Miller 2008). History has shown that new technologies provide effective tools for aspiring politicians: Roosevelt used the radio, John F. Kennedy the TV, Obama the Internet, and Trump heavily relied on social media, primarily Twitter, to promote his messages directly without any intervention by established gate keepers.[2] What all of these presidents did was using disruptive technologies (a) and developed strategies to benefit from these new communication patterns (b) in order to influence citizens. To further utilise behavioural phenomena (c), creating a diversified, fractured media landscape is highly advantageous. A fractured media system composed of binary subsystems that either support or oppose political and societal movements is prone to manipulation and tampering. Adducting Luhmann's logic of binary-coded systems, such subsystems would work according to partisan/non-partisan, liberal/non-liberal, mainstream media/non-mainstream media et cetera, thereby showing true colours and facilitating political leverage and influence. In addition, these subsystems in a fractured media landscape are perfect breeding grounds for alternative views, lies and propaganda; on top of that, surrounding communities are easily manipulable.

Analysing the 1988, 1992 and 1996 US presidential elections, Watts et al. (1999) suggest:

> [C]riticisms of media coverage–driven by conservative elites, and reported and discussed in new stories during these campaigns–have been a substantial influence behind the rising public perception of liberal media bias.

[2]This direct wire to the people lead to the surge of fact-checking projects as described initially.

> Indeed, analysis of a wide range of news content indicates that discourse on media bias in the past three presidential campaigns has focused overwhelmingly on allegations of liberal bias; has emanated increasingly from political elites (particularly if the claim is of liberal bias); and has predominantly conceived of the bias as existing across the entire media industry. (Watts et al. 1999, p. 166)

In a similar way Ladd (2010, p. 42) concludes that his study on antipathy towards news media "indicates that elite rhetoric can also powerfully influence the public's attitudes toward the institutional news media." When discussing the role of elites within this field of political communication, Stroud (2014) addresses "*selective production*", a phenomenon where elites are more open to giving interviews to agnate media. Recent incidents after the White House allegedly excluded certain members of the White House press corps due to their affiliation with critical or liberal news outlets (Mullin 2017), certainly substantiate this claim. Sykes (2016), who quit hosting a conservative political talk show, stated that the "conservative media is broken" and after years of efforts the conservative media "had succeeded in persuading our audiences to ignore and discount *any* information from the mainstream media. Over time, we'd succeeded in delegitimizing the media altogether – all the normal guideposts were down, the referees discredited". Before this op-ed for *The New York Times* he gave an interview to the *Business Insider* considering the conservative media who "basically eliminated any of the referees, the gatekeepers" (Darcy 2016) as responsible for disinformation, misinformation and propaganda proclaimed by Trump—partisan news outlets basically serve as mouthpieces for political elites.

Within this context, times for journalism seem dire. Besides these external influences as described above, factual journalism and in particular data-driven journalism has feared for its integrity during recent years. While fact-checking sites have been correcting claims or debunking lies, many legacy media outlets chose partisanship over objectivity, giving the floor to political fights, undermining their own impartiality. Data-driven journalism, seeking objectivity through forecasting based on statistical models and deriving facts from numbers, has been trying to evade this vicious circle of being instrumentalised and subsumed. In theory, journalism based on statistics, data collected by governmental bodies, academia or other research institutions and reported by professional

journalists trained in data analysis and conveying numbers in the form of widely comprehensible journalistic products, sounds like salvation for skewed media systems. It is not. Kayser-Bril (2016) states that "Truth is a niche market" as news outlets would rather bow to interests of donors, advertisers or readers' interests without regard to quality and fact-based journalism. He continues that some US outlets such as Vox or FiveThirtyEight would very successfully serve this niche market. During 2016, however, fact-based journalism had to suffer massive setbacks. After most analytical and explanatory outlets misinterpreted—or at least reported their findings rather ingenuously—and ultimately wrongfully forecasted the results of US elections and Brexit referendum, they saw themselves confronted with accusations: Their methodology would be flawed, even impractical as a journalistic practice. Data journalists and statisticians were accused of blindly applying statistical models while losing touch with reality; at last, due to their painfully wrong forecasts, Gabler (2017), among others, concluded that "[n]o one will ever believe Nate Silver or Nate Cohn or The Huffington Post aggregator or any other polling data ever again. Political Moneyball failed spectacularly".

Leading poll analyst Nate Silver repeatedly warned that chances of a different outcome of the US election than expected are not that small (Silver 2016a, b; Milbank 2016)—in fact, his team gave Trump the highest chances of winning across all forecasts. Even if certain results are more unlikely than others, this does not mean that they are impossible. After Trump's nomination, that analysts—again—did not see coming, *The New York Times*'s Rutenberg (2016) called out data journalism and his ex-colleague Silver and demanded a return to shoe-leather reporting, "to get a good night's sleep, and then talk to some voters". At this point we need to paint a more nuanced picture of data journalism: Forecasts based on advanced statistical modelling and likelihoods have a margin of error. Such projections have been enjoying enormous popularity over the past years, particularly as they claim to predict results of elections or referenda. After these failing predictions, negative repercussions do not only harm predictive data journalism but data journalism in general. Descriptive, explanatory and analytical data-driven journalism, that reports demographic or economic data, portrays communities and societies, and presents correlations of factors by connecting indicators, applies more basic statistical calculations. Clearly, we need to distinguish these two streams—predictive and descriptive data journalism.

Unfortunately, infighting will lead to a further disrupted media system. Instead of working together, trying to right the wrongs and developing better statistical models or even bundling forces of traditional reporters and data journalists, different journalistic streams appear to resort to pointing fingers. These frictions hurt the integrity of quality journalism as well as damage the reputation of data journalism—and data journalism is an integral part of objective reporting in the fight against disinformation: "The declining authority of statistics – and the experts who analyse them – is at the heart of the crisis that has become known as 'post-truth' politics" (Davies 2017). It is important to acknowledge the value data-driven practices bear for journalism. A return to shoe-leather reporting would have journalists to rely on old-established sources such as press releases and official statements. In a post-truth era, these sources are likely to be peppered with malinformation and disinformation. Data sources—be it that they are provided by official governmental bodies, non-governmental organisations or whistleblowers—allow citizens and journalists to compare claims and statements with actual underlying data, ultimately holding elites and governments accountable. As Segnini (2017) puts it: "Data can help a reporter fill in the missing pieces, and it also allows journalists to expose a universe beyond the one sources are willing to share." In that sense, citizens should look to journalists rather than state officials when it comes to constructing models of reality.

Certainly, aspects of shoe-leather reporting need to be inherent to data journalism. *The New York Times* alumni Sarah Cohen says that "if a data reporter doesn't go out into the field and report, they are only doing a third of their job" (Pierce 2015). Data journalism could greatly benefit from enhancing stories and findings drawn from quantitative data with a qualitative narrative. This would mean appending personal accounts with rows of data. Giving these numbers—that once were abstracted from individual cases—voices and faces would make data journalism more tangible and relatable. A combination of quantitative and qualitative approaches would also invalidate the argument that data journalism exclusively targets data-savvy elites, requiring knowledge of statistics, the willingness to read numbers and digest complex data visualisations. Dickinson (2013) asks if data journalism is not "just creating a small, equally uncountable, data elite? Is it really just a good way to reposition (consolidate) journalism as gatekeepers?" Indeed, data journalism might run the risk of building an elitist circle of journalists who know statistics and computational

techniques as well as appealing to an audience that is already data-oriented while neglecting the majority of readers. To attract a wider audience, data journalism needs to fulfil two tasks: First, educating their readers. Just as the many journalists who got into data, they need to pass their knowledge on to their audience. By explaining their calculations, their modelling, used statistical indicators et cetera, journalists do not only create transparency regarding their employed methods, but they could also dispel fears of seemingly over-complicated stories. In this regard, data journalists need to be more than data reporters: They have to become educators. Secondly, within a holistic approach, data journalism needs to develop appealing and intelligible storytelling formats. Translating data and statistics into comprehensible journalistic products does not equal long reads packed with numbers and dozens of interactives. Accessible storytelling means keeping in mind the reader's habits and socialisation. Alvin Chang of Vox considers a lack of context, particularly when covering complex topics, as problematic when trying to connect with the reader: Laying out the context not only makes a story more comprehensible by describing the background, the underlying mechanics of a problem, but he also regards it as a means to fight disinformation. On another note, Chang argues that a more visual journalism can create empathy and lead to more inclusivity in contrast to fabricating exclusivity through the use of jargon (Adams 2017).

If data journalism wants to constitute itself as an objective corrective in mainstream media and play a major role in today's news dissemination and news consumption, it needs to continue developing its profile as a data-driven and evidence-based form of journalism. Whether or not it succeeds in doing so, depends on many interfering variables as laid out in this essay. An objective and scientific analysis and interpretation of data can promote a journalism culture that combats disinformation and misinformation, and also holds accountable authorities within digital public spheres.

References

Adair, B. (2014, April 4). *Duke Study Finds Fact-Checking Growing Around the World*. Retrieved from https://reporterslab.org/duke-study-finds-fact-checking-growing-around-the-world/.

Adams, J. (2017, November 14). Q&A: Vox Data Guru on How Cartoons Can Help Simplify Complex Issues. *Columbia Journalism Review*. Retrieved from http://www.cjr.com.

Brock, T. C., & Balloun, J. L. (1967). Behavioral Receptivity to Dissonant Information. *Journal of Personality and Social Psychology, 6*, 413–428. https://doi.org/10.1037/h0021225.

CrossCheck. (2017). CrossCheck. *A Collaborative Journalism Project*. Retrieved from https://crosscheck.firstdraftnews.org/france-en/.

Darcy, O. (2016, August 26). Donald Trump Broke the Conservative Media. *Business Insider*. Retrieved from http://www.businessinsider.com.

Davies, W. (2017, January 19). How Statistics Lost Their Power—And Why We Should Fear What Comes Next. *The Guardian*. Retrieved from http://www.theguardian.com.

Dickinson, A. (2013, August 13). *Does Data Journalism Help Democracy?* Retrieved from http://digidickinson.net/2013/08/13/does-data-journalism-help-democracy/.

Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Evanston, IL: Row, Peterson.

Gabler, N. (2017, January 24). Five Ways the Media Bungled the Election. *Columbia Journalism Review*. Retrieved from http://www.cjr.org.

Graves, L. (2018). Boundaries Not Drawn: Mapping the Institutional Roots of the Global Fact-Checking Movement. *Journalism Studies, 19*(5). https://doi.org/10.1080/1461670x.2016.1196602.

Kayser-Bril, N. (2016, September 26). *Data-Driven Journalism in the Post-truth Public Sphere*. Retrieved from http://blog.nkb.fr/datajournalism-in-the-posth-truth-public-sphere.

Klapper, J. T. (1960). *The Effects of Mass Communication*. Glencoe: Free Press.

Ladd, J. M. (2010). The Neglected Power of Elite Opinion Leadership to Produce Antipathy Toward the News Media: Evidence from a Survey Experiment. *Political Behavior*, *32*(1), 29–50. https://doi.org/10.1007/s11109-009-9097-x.

Lewin, K. (1947). Frontiers in Group Dynamics: II. Channels of Group Life; Social Planning and Action Research. *Human Relations*, *1*(2), 143–153. https://doi.org/10.1177/001872674700100201.

Milbank, D. (2016, November 8). No Matter Who Wins the Presidential Election, Nate Silver Was Right. *The Washington Post*. Retrieved from http://www.washingtonpost.com.

Miller, C. C. (2008, November 7). How Obama's Internet Campaign Changed Politics. *Bits/The New York Times*. Retrieved from http://bits.blogs.nytimes.com.

Mullin, B. (2017, February 24). White House Press Corps Rebels as Colleagues Are Excluded from Gaggle. *Poynter*. Retrieved from http://www.poynter.org.

Pierce, E. (2015, September 29). New York Times Editor: Data Journalism Starts with People. *Downtown Devil*. Retrieved from http://www.downtown-devil.com.

Rogers, J. F. (2015, February 13). *Are Conspiracy Theories for (Political) Losers?* Retrieved from https://yougov.co.uk/news/2015/02/13/are-conspiracy-theories-political-losers/.

Rutenberg, J. (2016, May 5). The Republican Horse Race Is Over, and Journalism Lost. *The New York Times*. Retrieved from http://www.nytimes.com.

Ryssdal, K. (2016, October 13). *Poll Finds Americans' Economic Anxiety Reaches New High*. Retrieved from https://www.marketplace.org/2016/10/13/economy/americans-economic-anxiety-has-reached-new-high.

Segnini, G. (2017, August 3). Data Empowers Journalism Independence in Trump's Era. *Columbia Journalism Review*. Retrieved from http://www.cjr.com.

Silver, N. (2016a, November 1). Election Update: Yes, Donald Trump Has a Path to Victory. *FiveThirtyEight*. Retrieved from http://www.fivethirtyeight.com.

Silver, N. (2016b, November 11). Why FiveThirtyEight Gave Trump a Better Chance Than Almost Anyone Else. *FiveThirtyEight*. Retrieved from http://www.fivethirtyeight.com.

Stencel, M. (2017, June 30). *Fact-Checking Booms as Numbers Grow by 20 Percent*. Retrieved from https://reporterslab.org/fact-checking-booms-as-numbers-grow-by-20-percent/.

Stroud, N. J. (2014). Selective Exposure Theories. In K. Kenski & K. H. Jameson (Eds.), *Oxford Handbook of Political Communication*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199793471.013.009.

Sykes, C. J. (2016, December 15). Charlie Sykes in Where the Right Went Wrong. *The New York Times*. Retrieved from http://www.nytimes.com.

Valverde, M. (2017, April 10). *Donald Trump Changes Yardstick in Claim About Southern Border Apprehensions*. Retrieved from http://www.politifact.com/truth-o-meter/statements/2017/apr/10/donald-trump/trump-illegal-immigration-down-64-percent-march-tr/.

Watts, M. D., Domke, D., Shah, D. V., & Fan, D. P. (1999). Elite Cues and Media Bias in Presidential Campaigns: Explaining Public Perceptions of a Liberal Press. *Communication Research, 26*(2), 144–175.

White, D. M. (1950). The "Gate Keeper": A Case Study in the Selection of News. *Journalism & Mass Communication Quarterly*, *27*(4), 383–390. https://doi.org/10.1177/107769905002700403.

Check for updates

CHAPTER 23

The Future of Investigative Journalism in an Era of Surveillance and Digital Privacy Erosion

Julie Posetti

Introduction

Many investigative journalists who are struggling to uphold their ethical commitment to protect their sources in the digital era are changing their practices significantly in the aftermath of Edward Snowden's explosive 2013 revelations (Harding 2014). "Going back to analogue basics" is the new normal when dealing with confidential sources. Other methods of defending source protection in the digital era include increasing reliance on encryption, outsourcing the risks to confidential information clearing houses like the International Center for Investigative Journalists (ICIJ), implementing digital security training within newsrooms, and journalists training their sources in secure methods of communication. This chapter builds on a 2017 study on the state of the legal and normative frameworks supporting confidential source protection in 121 countries:

J. Posetti (✉)
Reuters Institute for the Study of Journalism, University of Oxford, Oxford, UK

© The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1_23

Protecting Journalism Sources in the Digital Age (Posetti 2017). The study was produced for UNESCO by the author.[1]

Background to a Crisis

The 2017 UNESCO study identified worrying global trends: Journalists' sources are increasingly being caught up in surveillance nets, and laws protecting journalists from being forced to reveal their sources are under significant threat of erosion, restriction and compromise.

It used to be possible to promise confidentiality to sources—guaranteeing protection of their identities, even on pain of jail—in countries where legal source protection frameworks were robust. But the UNESCO study found that during the decade from 2007 ethical commitments to and legal protections for confidential sources have been increasingly:

- Eroded by national security and anti-terrorism legislation
- Undercut by surveillance—both mass and targeted
- Jeopardised by mandatory data retention policies and pressure applied to third party intermediaries (like ISPs, telcos, search engines, social media platforms) to release data
- Outdated when it comes to regulating the collection and use of digital data gathered by journalists

Dozens of international journalists, editors, lawyers and freedom of expression experts were interviewed for the study which concluded that journalists were taking measures to adapt their practices in an effort to shield sources from increasing risks of exposure. But threats to anonymity and encryption undermine these adaptations. The results include more reluctant sources, more tedious reporting strategies, and the risk that much public interest journalism will be hampered. So, the attention of investigative journalists and their editors is necessarily turning to risk assessment, self-protection and source education.

[1]Note: The author produced the UNESCO study under contract to WAN-IFRA and the World Editors' Forum (Paris) where she was based as Research Fellow and Editor. Disclaimer: This body of work is the responsibility of the Author, and the ideas and opinions expressed in it are not those of UNESCO and do not commit the Organisation.

How Much Confidence Do Investigative Journalists Have in Their Ability to Protect Their Sources?

At the time of our interview in his London office in late January 2015, outgoing Editor-in-Chief of The Guardian Alan Rusbridger was despondent about the threat to investigative journalism posed by the erosion of source protection. "Well, I'm very gloomy," he said. "The limitations on existing legal frameworks supporting source protection in the UK are coming thick and fast. It's like fighting a 'Zombie War'," he said, waving his hands in exasperation.[2]

Rusbridger previously suggested that investigative journalism might not be possible in the post-Snowden era. That concern was echoed by the Committee to Protect Journalists' Advocacy Director Courtney Radsch: "I think that we are really potentially looking at an environment where it becomes virtually impossible for journalists to protect their sources—where journalists are no longer even needed in that equation, given governments' broad surveillance powers," she said.

Bolivian investigative journalist Ricardo Aguilar expressed serious concern about the reliability of legal source protection. He was charged with espionage and faced up to 30 years jail after refusing to reveal his source on a 2014 *La Razon* story. "…[M]ass surveillance, data retention and the appeal of [the] National Security category leaves the protection of secret sources in latent vulnerability," he said.

Director of the US-based International Consortium of Investigative Journalists (ICIJ), Gerard Ryle, was similarly direct. "I'm not confident that there is any protection at all to be frank. … I would say as a general rule these days, much more than in the past, it's very difficult to protect sources because of the fact that electronic communications can be backtracked and people can be found much easier than they may have been in the past," he said. Ryle, who oversaw the global investigative journalism projects Panama Papers (ICIJ 2016), Offshore Leaks (ICIJ 2013), Luxembourg Leaks (ICIJ 2014) and other major investigations, once faced the threat of jail in Australia while reporting on police corruption for *The Age*, after refusing to give up a source to an ombudsman's inquiry.

[2]Note: Unless otherwise indicated, all quotes featured in this chapter are drawn from research interviews conducted for *UNESCO's Protecting Journalism Sources in the Digital Age* (2017).

In Sweden, where source protection legislation is so strong that journalists can be jailed for revealing their confidential sources, top investigative journalists are taking extraordinary measures to protect them from the impacts of mass surveillance, and other risks. One of the threats identified by the director of the investigative unit at Sweden's national public radio (Sveriges Radio), Fredrik Laurin, is that of police seizing digital content due to gaps in source protection legislation in his country[3]: "…It's not an exception—this is definitely the modus operandi. The police, they don't go into newsrooms very often here, but when they do they have no problem in grabbing digitally stored information."

The Chilling Effect

Co-founder of Pakistan's Centre for Investigative Reporting, Umar Cheema, believes his status guarantees that he is under surveillance and his sources know it. "I am a prominent journalist, a distinction with its own advantages and disadvantages. Some [sources] tend to approach me out of respect and belief that I am the right person to be taken into confidence. Others hesitate, fearing any contact with me will put them on [the] radar screen since I am under surveillance, from phone to emails, and [my] social media accounts are monitored."

Cheema was kidnapped and tortured in 2010. In the course of his captivity, his sources were compromised. "The captors, who I strongly suspect belonged to our premier intelligence agency, took away my mobile phone, apparently for investigating in detail about my professional contacts through my phone contacts," he said. "Some of my sources, who had shared information about national security, were coerced into silence. They never contacted me afterwards, other than telling me … about the harassment they had to face." Cheema said that threats to his safety sent via phone and email are now routine.

International Editor of Algeria's *El Watan* newspaper, Zine Cherfaoui, said that sources now increasingly require face-to-face meetings. "Since Snowden and mass surveillance, sources speak with difficulty and people don't have as much confidence. To really discuss with people we prefer to avoid electronic means or social networks. The Snowden Affair turned upside down the work of journalists. … It's harder to speak

[3]See Thematic Study 2 in *Protecting Journalism Sources in the Digital Age* (Posetti 2017).

to people. We really have to go out and meet them. It's face to face," Cherfaoui said.

However, it should be noted that the risk of exposure travels with journalists heading to face-to-face meetings with sources if the route they take is subject to security camera surveillance, or they travel with traceable mobile devices that deliver geolocation data.

The cost of digital security technology, training and legal fees connected to source protection in the post-Snowden era (Bell and Owen 2017) also represents a significant chilling effect on investigative journalism. In 2015, *The Guardian* was spending about a million pounds more a year on legal fees than they did five years earlier, according to Rusbridger: "It's definitely having a bad effect on the overall ability to report," he said. "Of course, once you get into secure reporting there is a significant cost … in trying to create a safe environment where we feel we can offer our sources the kind of protection that they deserve." Rusbridger pointed to the devastating impact of the changed landscape on regional newspapers, in particular. "[They] can't afford to get tied up in defending their staff, or equipment, or the IT," he said.

But Is This Not a Golden Age for Investigative Journalism?

"Technology is allowing information to be leaked on a vast scale. … For me as a journalist we're in boom times, because you're able to get information that's incredibly detailed and you're able to get stories that you couldn't possibly [get before]," ICIJ's Gerard Ryle said, declaring the digital era a "golden age for journalism," despite the risks.

Prominent Jordanian investigative journalist and founder of the Arabic Media Internet Network, Daoud Kuttab, echoed Ryle's view of the digital era: "On the one hand I think it has accelerated and widened the amount of data available to everyone and made it very easy to transfer information and documents. Now you can put thousands of documents on a USB so you don't have the problem of having to carry things out of offices—you can email, send as an attachment. But at the same time governments are able to invade your privacy much easier and get information."

Editor-in-Chief of Argentina's *La Nacion*, Carlos Guyot, also acknowledged the significant benefits of digital era investigative reporting involving confidential sources, including access to leaked documents that would have been impossible to get even five or ten years ago. "New

technologies bring new challenges with them, but also new opportunities, like encrypted conversations via new software, although this must be combined with old fashioned practices. … There is nothing like a face-to-face meeting with a source," he said.

However, one of the risks of this data boon is the rush to legislate against the impacts of leaks, according to Gerard Ryle. "The leaks are getting bigger, therefore the law is scrambling to catch up…and that's the danger for authorities, and for people who want secrecy, and I think that there is a push generally across the world to try and cope with this," Ryle said. "[It's] a problem for governments, agencies, any organisation that wants to keep secrets. It's becoming more and more difficult to keep those secrets."

Just Assume You Are Being Watched

How do reporters protect their confidential communications with sources in the age of surveillance? "I'm more careful with any digital platform that I'm involved in—whether it's email, phone or any other digital format. I assume that [I am] probably being watched, listened to, or read. That's my starting point and I take it from there," Daoud Kuttab said.

ICIJ's Gerard Ryle adopts the same mode. "Don't put things in writing, don't do certain things if you don't want them to come out afterwards. You have to assume that everything you do is being recorded or traced," he said.

A change of practice in managing digital communications is required in response—at both the personal and professional levels—according to Deputy Director of the Tow Centre for Digital Journalism, Susan McGregor. "It means that we have to be thoughtful about our devices and our communications in the way that most of us aren't accustomed to doing yet. … Some of the habits we've developed as private individuals, taking our phone everywhere, always having wifi on, emailing everything, we're just going to have to think differently about those things when it comes to working with sources," she said.

Going Back to Analogue Basics

"I know investigative journalism happened before the invention of the phone, so I think maybe literally we're going back to that age, when

the only safe thing is face-to-face contact, brown envelopes, meetings in parks or whatever," Alan Rusbridger said, despairing that investigative journalism based on confidential sources may not be possible in the digital age, unless journalists go 'back to (analogue era) basics'.

UK QC Gavin Millar, who has advised *The Guardian*, tells his clients to revert to traditional methods of investigative journalism. "They actually have a contract phone and throw it into the Thames at the end of each week, they will meet sources in pubs, write notes, hide the notes… in distant places where people can't get them if their houses are searched by police, and some of them are very, very good at it," Millar said.

Bolivia's Ricardo Aguilar avoids using digital communication in order to protect his sources. "Extreme distrust is the only defence against the possibility of a raking of secret sources in email accounts or social networks," he said. And *La Nacion's* Carlos Guyot said his investigative journalists were spending a lot more time on the road now. "Our main investigative reporter drove for three hours to a different city for a 15-minute conversation with a source, and drove back to our newsroom. If we are willing to endure the challenges, we can still do good journalism," he said.

According to *El Watan's* Zine Cherfaoui, journalists in the Middle East and North Africa have also reverted to face-to-face meetings with confidential sources, being particularly concerned about email communication. "We've become very cautious with social networks and everything that is electronic. Generally, we prefer to meet the source in person when it is very important. … Because of mass surveillance and new anti-terrorism laws we like to avoid social networks."

Swedish Lawyer and Press Ombudsman Pär Trehörning agrees: "I've talked to a lot of editors and the best thing to do today is to write an ordinary letter." Three journalists interviewed for the UNESCO study mentioned the trend of relying on chat apps as a more secure form of source interaction than email, but Mexican journalism safety expert Javier Garza Ramos warned against such an approach. "If we're sloppy and we say everything we know about our sources on our WhatsApp, then of course the government is going to find out who our sources are, or whoever is spying on us," he said.[4]

[4]See also: Ramos Garza, J. (2016). *Journalist Security in a Digital World: A Survey.* Available from CIMA website: http://www.cima.ned.org/wp-content/uploads/2016/03/CIMA-Journalist-Digital-Tools-03-01-15.pdf.

Simple approaches like stretching the timeline between contact with a source and publication of their leaks can also be used to disguise connections and minimise the chance a source will be "caught", ICIJ's Gerard Ryle said. "I mean the more layers you can put between you and the source the better, and a lot of that is time. … If someone gives you some really hot information, the temptation is to publish that right away, but that's also when your source is potentially at most risk."

Taking Responsibility for Digital Security

In 2018, it is not just investigative journalists and war correspondents who need to deal with digital era threats to source protection. According to Alan Rusbridger, the net is widening. For example: "It's become increasingly hard to report on the national health service because you know they all have confidentiality agreements, so if you're a health reporter you probably want to make sure that you begin to understand this stuff."

The other factor to consider is that seemingly innocuous local stories built on anonymous sources can turn into large-scale investigative journalism projects. From little stories big stories grow. But careless initial contact with a source makes such a person increasingly vulnerable as the story develops. Swedish public radio's Fredrik Laurin said journalists are underdeveloped when it comes to protecting sources in the 'digital hemisphere'. "Very few journalists use encryption and know how to use it—it's not in their toolbox and that is a major problem," he said. "And when you do come into contact with sources…you often get confronted with very important questions—how do you, in reality, protect this source? Are you going to store the information on the company server? How are we going to communicate? What level of encryption do you use? Serious questions." According to Laurin, his team's digital security expertise gives them an edge in journalism based on confidential sources.

Laurin's hardcore dedication to digital security in the interests of protecting his sources may seem extreme, but it needs to be understood in the context of the Swedish legal source protection framework that actually criminalises unauthorised source revelation. "It's me, Fredrik who goes to prison if you are my source and I lose my notebook at the bar and your name comes out because of that. That's my fault and I go to prison. That's why I don't use…Facebook," he said. Laurin also bans his staff from using certain devices due to concerns about security

vulnerabilities. “I need to survey—which I do, very thoroughly—who my suppliers are. I know exactly where my server is standing, I know exactly what the contract says, the hard discs in that server are named in my name, with my phone number. There’s a tag on the material that says this material is protected according to the Swedish constitution.”

However, ICIJ’s Ryle, who remains optimistic about the future of investigative journalism in the digital age despite the threats to source protection that he acknowledged, said that too many journalists are growing unnecessarily paranoid. “There are some reporters I know (who are) completely paranoid about their computers—they’re fantastic at encryption, everything is offline. But so what? Most of what they’re working on isn’t relevant.”

Another issue to consider: Digital security measures designed to protect sources can be unwieldy and time-consuming, and these factors remain a deterrent to many investigative journalists. “When we were doing the Offshore Leaks project we started off by trying to encrypt a whole email communication with everyone we were working with, it became a complete nightmare, because, first of all not all of us are very technological, including myself, and it became a hindrance to communication,” Ryle said.

However, encryption was the default requirement of ICIJ during the Panama Papers investigation which saw 400 journalists from 80 countries working collaboratively for 12 months to produce high-impact journalism that ultimately brought down a government. To participate in the project, all members were required by ICIJ to use two-factor authentication and encryption. Prior to the Panama Papers investigation, a survey of 118 ICIJ members found that 47% of journalists were mostly unaware of PGP encryption or two-factor authentication, and 45% had never used them (McGregor et al. 2017). However, after the Panama Papers collaboration, which used a secure interactive platform known as the Global I-Hub, over 60% of surveyed journalists reported to researchers that the encryption requirements for the collaboration, which demanded all documents and messages (no matter how sensitive) be encrypted, were “easy” to use (McGregor et al. 2017).

Recognising the role of encryption in the defence of journalism based on confidential sources, the UN Special Rapporteur for the Promotion and Protection of the Right to Opinion and Freedom of Expression, David Kaye, has declared encryption an important tool to secure human rights as a range of States seek to limit the availability

of encryption and ensure 'backdoor access' to encrypted data (UN Human Rights Council 2015).

Journalists Need Training in Digital Security, But so Do Their Sources

There is a new trend emerging in reference to source protection: journalists are beginning to train their sources in digital security to help them ensure their anonymity. *La Nacion*'s Carlos Guyot said: "If we want journalism to survive and flourish in the 21st century, there is no other option than give our reporters, and sources, the tools necessary to do their jobs."

Alan Rusbridger acknowledged this challenge. "But because often sources are of interest to people with access to surveillance equipment, corporate or government, it feels like an unequal battle really."

However, as Executive Director of Arab Reporters for Investigative Journalism (ARIJ) Rana Sabbagh pointed out, even the best training cannot keep up with global intelligence services: "[W]e train our journalists in encryption and how to protect their data, and tell them to always assume that everything you're doing online is accessible, because even if you give them the best software and training, the intelligence agencies are always a step ahead. They are using the latest technologies to decrypt the content, they are using technologies coming from countries that are supposed to protect free speech like the US and Switzerland."

Nevertheless, encryption may buy time in the course of an investigation, and it may at least keep other potentially hostile actors at bay—even if not the intelligence agencies.

Outsourcing Source Protection

In its global investigations that involve myriad international publishing partners, ICIJ essentially becomes the source: "By taking all the responsibility of source protection and also putting the responsibility on each organisation to do whatever it is according to their own laws. So we don't take responsibility for the publication of our projects in each country, each organisation has to do that, but in terms of giving them the information, we become the source. … In other words we give them the documents. … ICIJ is the source of the material," Gerard Ryle said.

Meanwhile, international news organisations have begun using secure drop boxes and collaborating on platforms designed to securely receive digital information from confidential sources. AfriLeaks, for example, is a Pan-African project that uses a highly secure mailbox designed to receive leaked documents, which connects investigative media houses to whistleblowers. It is operated by the African Network of Centers for Investigative Reporting. In Mexico, Mexicoleaks launched in 2015 with a similar mission.

Sourcesûre and Balkanleaks are Francophone and Bulgarian websites that follow the same model, allowing whistleblowers to upload secret documents anonymously. Sourcesûre, which is based in Belgium to take advantage of strong source protection laws there, was jointly established in February 2015 by France's *Le Monde*, and Belgian publications *La Libre Belgique*, *Le Soir de Bruxelles* and RTBF (Radio Télévision Belge Francophone). Yves Eudes, Sourcesûre's cofounder and a journalist at *Le Monde*, said that the cross-border, multi-platform collaboration between leading Francophone news organisations is a "spring of immunity" against coercion for journalists and their sources. "Unity is strength. This initiative could not have been launched by *Le Monde* or RTBF alone. Sourcesûre is underpinned by a whole spectrum of collaborators, from liberal to conservative media outlets, united by common journalistic values," he said.

Ultimately, is it sustainable to promise confidentiality to sources in an era when it is so easy to identify a source without the involvement of the journalist, especially considering it can be a life or death matter? ARIJ's Rana Sabbagh was clear in her response: "Even in the best and most democratic of countries, one can't promise that anymore. There is no 100 percent guarantee."

Conclusion

The erosion of legal frameworks designed to protect confidential journalistic communications internationally poses a serious threat to the sustainability of investigative journalism. The digital era threats undermining these protections include increasing surveillance, data retention policies, the handover of data by third party intermediaries, national security and anti-terrorism legislation overreach, and moves by States to ban or stymie encryption and anonymity. In response, journalists are increasingly

changing their investigative reporting techniques and reverting to analogue era practices and processes.

With the great power of digital data comes even greater responsibility to sources, and amplified risks that all journalists should be aware of. But the defence of investigative journalism based on information from whistleblowers and confidential sources requires a collective commitment to upholding the fundamental and interlinked human rights to privacy and freedom of expression. Civil society organisations, human rights defenders and States committed to the flourish of democracy must work cooperatively with journalists to limit the impacts of the growing global culture of surveillance and interception, and to ensure the sustainability of investigative journalism. This collaboration needs to focus on educating journalists, whistleblowers and confidential sources about the need for risk awareness and threat assessments, as well as equipping them with the technology necessary to undertake secure digital communications. It should also involve a commitment to advocate for broad rights to anonymity and encryption, and educate societies about the role and value of investigative journalism.

Seven Tips for Journalists to Help Make Communications with Sources and Whistleblowers More Secure

- Do not grow unnecessarily paranoid—instead act smarter, get properly equipped and go back to analogue-era basics where necessary
- Be aware that even face-to-face meetings can be compromised by the presence of geolocatable mobile devices and security cameras
- Assume you are being watched
- Encrypt your data
- While recognising the value of encryption and anonymity, be aware that using Tor, PGP and other forms of data encryption can 'red flag' digital communications with sources, i.e. such practices can make you and your sources a bigger target
- Recognise your ethical responsibility to protect your sources and consider training your confidential sources in digital safety and security
- Recognise that it may no longer be possible to guarantee protection for your confidential sources and consider the ethical implications of that realisation

References

Bell, E., & Owen, T. (Eds.). (2017). *Journalism After Snowden: The Future of the Free Press in the Surveillance State*. New York: Columbia University Press.

Harding, L. (2014). *The Snowden Files: The Inside Story of the World's Most Wanted Man*. New York: Vintage Books.

ICIJ. (2013). *Bahamas' Secrets Revealed*. https://www.icij.org/investigations/offshore/.

ICIJ. (2014). *Luxemburg Leaks*. https://www.icij.org/investigations/luxembourg-leaks/.

ICIJ. (2016). *The Panama Papers*. https://panamapapers.icij.org/.

McGregor, S., Watkins, E. A., Al-Ameen M. N., Caine, K., & Roesner, F. (2017). When the Weakest Link Is Strong: Secure Collaboration in the Case of the Panama Papers. In *USENIX Association, Proceedings from the 26th USENIX Security Summit* (pp. 505–522). Vancouver, Canada. Retrieved from https://www.usenix.org/system/files/conference/usenixsecurity17/sec17-mcgregor.pdf.

Posetti, J. (2017). *Protecting Journalism Sources in the Digital Age*. Retrieved from UNESCO website http://unesdoc.unesco.org/images/0024/002480/248054E.pdf.

UN Human Rights Council. (2015, May 22). *Report of the Special Rapporteur on the Promotion and Protection of the Right to Freedom of Opinion and Expression, David Kaye, A/HRC/29/32*. Retrieved from https://daccess-ods.un.org/TMP/7816334.3667984.html.

Index

0-9
360°, 99, 100, 113, 225

A
accountability, 24, 143, 146, 149, 151, 207
advocacy, 68, 75, 76, 148, 251
algorithm, 25, 108, 181
Al Jazeera, 96
Amnesty International, 68, 74–76, 138, 145, 146
anonymity, 147, 150, 250, 258–260
app, 98, 99, 114, 116, 122, 143, 144, 146–149, 151, 186
artificial intelligence, 32, 87
authenticity, 5, 61, 145, 147, 173–176, 180, 189

B
BBC, 24, 26, 127, 131, 204, 207, 218
Bellingcat, 5, 145, 205, 210–212
bias, 35, 75, 167, 242, 243
big data, 4, 127
blog, 48, 210–212
Brexit, 121, 238, 244
BuzzFeed, 23–25, 31, 100, 226

C
chart, 130
civil society, 71, 149, 197, 199, 200, 239, 260
Coda Story, 5, 222
collaboration, 1, 56–59, 61–63, 126, 149, 167, 239, 257, 259, 260
computation, 1, 3, 44, 245
computational journalism, 1, 5, 127
computer-assisted reporting (CAR), 15, 43, 127
confidential sources, 249, 250, 252, 253, 255–257, 259, 260
credibility, 33, 34, 144, 146, 163, 166, 173, 174
Crimea, 197, 198, 200, 207, 208, 220
crowdsourcing, 89, 148, 203, 210
cybersecurity, 60

© The Editor(s) (if applicable) and The Author(s) 2018

O. Hahn and F. Stalph (eds.), *Digital Investigative Journalism*,
https://doi.org/10.1007/978-3-319-97283-1

D
data
  big. *See* big data
  data-driven, 4, 5, 16, 45, 48, 69, 127, 243–245
  journalism. *See* data journalism
  open. *See* open data
  set, 15, 31, 34, 37, 61, 127, 166
  source, 52, 74, 245
  structured. *See* structured data
  visualisation, 24, 124, 125, 127, 131, 133, 245
database, 3, 9–11, 14–16, 27, 33, 35, 37, 60, 61, 64, 69–71, 176, 181, 182
data-driven journalism. *See* data journalism
data journalism, 3–5, 9, 10, 16, 17, 19, 21–27, 43, 44, 68–70, 72, 127, 243–246
democracy, 164, 196, 197, 199, 219, 260
digital
  communication, 73, 254, 255, 260
  journalism, 4, 128, 254
digitalisation, 3, 89, 171. *See also* digitisation
digitisation, 3, 20, 193, 199
disinformation, 1, 160, 194, 196–198, 201, 225, 238, 243, 245, 246
drone, 36, 37, 91–101, 209
drone journalism, 91, 95, 98, 101

E
election, 1, 45, 109, 114, 150, 197, 238, 242, 244
encryption, 60, 62, 150, 249, 250, 256–260
ethical, 26, 27, 95, 96, 150, 161, 162, 168, 249, 250, 260
eyeWitness, 143–151, 188

F
Facebook, 63, 100, 118, 139, 145, 200, 208, 233, 234, 238, 256
fact-checking, 6, 23, 199, 201, 237–239, 242, 243
facts, 3, 44, 144, 145, 157, 196, 197, 200, 208, 221, 243
Fader, 106–109
fake news, 144, 151, 196, 197, 199–201
FiveThirtyEight, 4, 45–48, 50, 51, 244
foreign correspondent, 219, 222
forensic, 104, 145, 165, 167, 168, 179, 180, 183, 184, 188, 189
fraud, 2, 64, 71, 196
freedom of information, 14, 20, 24

G
gatekeeper, 146
geolocation, 137, 142, 186, 200, 253
Germany, 34, 59, 107, 109, 118, 141, 142, 171, 199
Global Navigation Satellite System (GLONASS), 98
Global Positioning System (GPS), 98, 147, 175, 186
Google, 64, 93, 116, 128–130, 137–142, 145, 148, 149, 175, 176, 187, 188, 200
GOP, 51, 52
graphics, 44, 46, 47, 50, 62, 63, 113–115, 118–120, 122–127, 129, 131, 132, 181
Guardian, The, 3, 4, 15, 24, 34, 61, 122, 132, 145, 167, 224, 226, 251, 253, 255

H
hacker, 184, 206, 209

human rights, 67–76, 143, 144, 146–151, 161, 163, 167, 168, 257, 258, 260
Human Rights Watch, 68, 74, 146

I
identification, 84, 186, 209, 234
identity, 83, 84, 147, 223, 229
immersive journalism, 103, 109
interactivity, 24, 25, 124–126, 128–133
international consortium of investigative journalists (ICIJ), 56–65, 84, 249, 251, 253, 254, 256–258
investigation, 2, 5, 20, 24–26, 32, 36, 56–60, 62–64, 69, 70, 73, 74, 86, 89, 90, 104, 129, 144, 162, 166, 167, 183, 184, 189, 207–210, 212, 257, 258
investigative journalism, 2, 3, 5, 17, 19, 23, 36, 39, 60, 67, 109, 144, 158, 236, 251, 253–260

J
Jihadist propaganda, 230, 231, 234, 235
JPEG, 174, 181, 183, 187
jpg. *See* JPEG

K
Klein, E., 48

L
leak, 3, 56, 60, 62, 63
Leonhardt, D., 46
longform, 120, 225, 226

M
machine learning, 87
machine-readable, 74, 87
Maidan, 198
malinformation, 245
manipulation, 3, 6, 149, 171, 172, 180, 183, 186–188, 197, 200, 238, 242
map. *See* mapping
mapping, 9, 13, 14, 33, 45, 47, 71, 72, 74, 87, 97–101, 128, 130, 137, 141, 142, 149, 239
metadata, 2, 143, 145, 147–150, 175, 183, 184, 186–189, 200
MH17, 138, 140, 210, 212
Migrants' Files, The, 33, 34, 37
misinformation, 1, 6, 145, 151, 197, 238, 239, 243, 246
mobile, 80, 98, 99, 105, 107, 129, 132, 150, 179, 252, 253, 260
Monde, Le, 63, 259
multimedia, 63, 104, 113–118, 120–122, 127, 179, 183, 188, 233

N
New York Times, The, 3–5, 10, 13–15, 24, 45, 46, 104, 120, 122, 123, 163, 243–245
non-governmental organisation (NGO), 33, 68–71, 73, 143, 144, 149

O
Obama, B.H., 242
Obamacare, 49–52
objectivity, 159, 163, 238, 243
Offshore Leaks, 64, 251, 257
online journalism, 113, 120, 124
open data, 21, 72, 75

open source, 98, 107, 181, 204–206, 209–213
Organized Crime and Corruption Reporting Project (OCCRP), 5, 71, 80, 84

P
Panama Papers, 17, 23, 56, 58, 59, 63, 87, 251, 257
Paradise Papers, 56–65
photography, 158–160, 162, 164–168, 171–174, 224
PolitiFact, 237, 238
post-truth, 6, 160, 164, 196, 239, 245
privacy, 27, 64, 94–97, 188, 253, 260
propaganda, 6, 144, 151, 194–201, 204, 205, 225, 229–236, 238, 241–243
ProPublica, 37–39
Pulitzer Prize, 43, 56, 120, 237
Putin, V.V., 86, 200

Q
Quartz, 4

R
reliability, 45, 94, 145–147, 251
Russia, 86, 120, 121, 197, 200, 204–208, 210, 211, 213, 224
Ryle, G., 57, 65, 251, 253, 254, 257, 258

S
security, 15, 16, 61, 62, 74, 148, 150, 180, 208, 210, 222, 226, 249–253, 255–260
sensor, 180
shoe-leather reporting, 3, 61, 244, 245
Silver, N., 45, 46, 48, 244
smartphone, 108, 113, 119, 120, 143, 145, 147, 150, 171
Snowden, E., 9, 61, 88, 249, 252
social media, 2, 5, 6, 16, 17, 22, 37, 89, 108, 125, 138, 140, 141, 144, 145, 149, 151, 188, 189, 196, 211, 224, 225, 230, 233, 238, 242, 250, 252
social network, 59, 204, 209
source, 3, 19, 22, 23, 26, 28, 44, 63, 71, 75, 81, 138, 139, 145, 147, 151, 167, 173–175, 187, 189, 200, 249–251, 254–256, 258, 259
source protection, 64, 249–253, 256–259
spatial, 91–93, 100
Spiegel, Der, 3, 15, 114, 116, 117, 122
statistics, 3, 4, 9, 13, 14, 25, 26, 35, 69, 70, 81, 118, 127, 199, 238, 243, 245, 246
StopFake, 145, 199, 200
storytelling, 5, 24, 27, 48, 90–92, 99–101, 104–107, 109, 118, 120, 121, 124, 127, 129, 131, 221, 223, 225, 226, 246
structured data, 33, 36
surveillance, 61, 62, 250–255, 258–260
sustainability, 259, 260
Syria, 16, 206, 207, 210, 219, 224

T
tampering, 3, 147, 171, 180, 187, 188, 196, 242
technological innovation, 94
transparency, 3, 71, 75, 207, 238, 246
Trump, D.J., 21, 24, 45, 56, 114, 237
trust, 56–59, 189, 199, 200, 212
Twitter, 1, 25, 118, 122, 145, 211, 212, 233, 238, 242

U
Ukraine, 72, 138, 140, 149, 194, 197–201, 204–212, 219, 220, 224
UNESCO, 250, 251, 255
United Kingdom (UK), 22, 23, 81, 143, 150, 166, 210, 231, 241, 251, 255
United Nations (UN), 33, 67, 69, 73, 146, 219, 257, 258
United States (US), 9, 10, 14, 15, 31, 34, 36, 37, 55, 120, 199, 207, 224, 237, 242
United States of America (USA). *See* United States (US)
Upshot, The, 45–47, 50, 51
user-generated, 144, 200, 213, 223

V
verification, 6, 90, 137, 144–146, 149, 151, 171–173, 179, 188, 189, 197, 200, 204, 210, 211, 213
virtual reality (VR), 99–101, 103–109, 120, 165
visual analytics, 4, 5, 89
visualisation, 1, 24, 37, 124, 125, 127, 128, 130, 131, 133
visual journalism, 95, 246
Vox, 4, 45, 48, 49, 51, 244, 246

W
Washington Post, The, 34, 48, 61
watchdog, 2
WhatsApp, 145, 234, 255
whistleblower, 61, 88
Wonkblog, 48

Y
YouTube, 105, 138–140, 145, 205, 233, 234