

Department Of Journalism Studies.

Data journalism & audience engagement in the post-pandemic era

27 June 2022, Sheffield

Welcome to the symposium

During the Covid-19 pandemic, the use of statistical figures and data visualisations in news coverage embedded data journalism into daily news reporting.

Data visualisations are thought to help effectively communicate with audiences and make it easier for them to understand government health messages and related news stories. During the pandemic, effectively sharing data was a crucial aspect of keeping the public informed.

Against this backdrop, critical questions arise about the current and future use of data in journalism and news reporting.

This symposium will explore the importance of data in journalism, examine its potential to engage audiences, reflect on the experiences gained and lessons learned during the pandemic, and discuss our shared vision for data journalism and audience engagement in the post-pandemic landscape.

We extend a warm welcome and our thanks to all our guests today, each of whom brings specialist knowledge of the use of data in journalism and the media.

Dr Jingrong Tong

Senior Lecturer in Digital News Cultures

Programme

Speakers and presenting authors are listed in blue

10:00-10:10 Opening remarks: Professor Jackie Harrison

10:10-11:40 Session 1: Data journalism during Covid-19

Chair: Jingrong Tong

Louis Goddard: Vaccinating data journalism: applying lessons from the Covid-19 pandemic to a new phase in the news cycle

Carmen Aguilar García: Sky News: Automating the Covid-19 story in the newsroom

Farzeen Heesambee: Data journalism and the Covid-19 pandemic: analysis of data dissemination strategies across broadcast, online and print news media

Tshuma Lungile: 'When two strangers meet': An uptake of data journalism by investigative journalists in Zimbabwe in the Covid-19 era

11:40-12:00 Break

12:00-13:30 Session 2: Data journalism and audience engagement Chair: Irini Katsirea

Neil Thurman, Florian Stalph & Sina Thaeslerkordonouri: Exploring audience perceptions of, and preferences for, data journalism

Jingrong Tong: UK audiences' perceptions of Covid-19 data visualisations: a case study

Jason Martin, Lindita Camaj & Gerry Lanosga: Data journalism and audience engagement in 34 countries: Instrumental or substantive?

Mathias Felipe de Lima Santos & Marília Gehrke: Credibility, objectivity, and transparency of data journalism during COVID-19: The audience's expectations when data leads to misleading information

13:30-14:20 Lunch

14:20-15.50 Session 3: Reflecting on the past and the present:
Implications for data journalism in the future
Chair: Claire Miller

Susan E McGregor & Joanna Kao: Supporting longhaul reporting: Data integration and management in the newsroom

Ross Lydall: Data journalism: do we need to get back to the office to do it properly?

Peter Sherlock: The BBC's Shared Data Unit: Data journalism and the public interest

Nathan Sparkes: "Does having a TV in your bedroom lead to more sex?", and more improbable headlines backed by questionable data: The ethics of data journalism

15:50-16.00 Closing remarks: Claire Miller



Carmen is an award-winning journalist specialised in data journalism. She is currently a senior data journalist at Sky News, where she uses a wide range of tools and techniques to find data-driven stories and tell them in compelling visual narratives. Before Sky News, Carmen worked as a TV and multimedia reporter in prominent media companies in Spain and Chile.

Carmen Aguilar García, Senior Data Journalist, Sky News



Lindita is a former journalist with a doctoral degree from Indiana University and a research agenda at the intersection of journalism, technology, politics and access to information. Her latest research explores citizen engagement with quality news in digital spaces and tries to identify best journalism practices that foster civic behavior. Her scholarly work has been awarded by the AEJMC and published in numerous high impact journals.

Lindita Camaj, Associate Professor & Director of Graduate Studies, Jack J. Valenti School of Communication, University of Houston



Marília holds a PhD. in Communication and Information from the Federal University of Rio Grande do Sul, where she defended her dissertation focused on data journalism and transparency. She is founder of the research group DataJor) and edited the media literacy hypermedia guide "To post or not to post?". Her research interests are digital journalism, data journalism, transparency, news sources and audiences, and disinformation studies.

Marília Gehrke, Post-doctoral Researcher, Trust and News Authenticity project, Digital Democracy Centre, University of Southern Denmark



Louis leads data-driven investigative work on a campaign against Russian fossil fuel exports. He finished a PhD in contemporary poetry at the University of Sussex in 2016 then spent two and a half years as a data journalist at The Times. He worked on over a dozen front-page stories, including flagship investigations into cyber-crime, tax and government.

Louis Goddard, Senior Data Investigations Adviser, Global Witness



Jackie has been with the University of Sheffield since 1996 and chair of the interdisciplinary research institute Centre for Freedom of the Media (CFOM) since 2008. In her role with UNESCO, she actively promotes in-depth collaborative academic research with global organisations. Jackie has acted as expert adviser to many UK government departments, the European Commission and multiple regulators and committees.

Jackie Harrison, Professor of Public Communication, University of Sheffield; UNESCO Chair on Media Freedom, Journalism Safety and the Issue of Impunity



Farzeen is on a studentship for an inter-disciplinary PhD specialising in data dissemination strategies in journalism during Covid-19. She is interested in cross-curricular research in social sciences and humanities, with qualifications in linguistics, education and postcolonial literature. She was a research assistant for the DFG-AHRC funded SpaTrEM project at Nottingham Trent University on bilingual magazines post World War Two.

Farzeen Heesambee, PhD Candidate, University of Derby



Joanna is the tech lead of the Financial Times' visual and data journalism team. She has a computer science degree from MIT and an MBA from IE Business School. She was previously a data visualization journalist at the FT's New York bureau, an interactive developer at Al Jazeera America, and taught data visualization and accessibility at Columbia Journalism School and Parsons School of Design.

Joanna Kao, Visual & Data Journalism Tech Lead, Financial Times



Before joining the university in 2015, Dr Irini Katsirea was Senior Lecturer in Law at Middlesex University. She served as Visiting Fellow at the Institute of Advanced Legal Studies and is a member of the Athens Bar. She holds a First Examination in Law from the Free University of Berlin, an LLM in European and International Trade Law from the University of Leicester and a PhD from Magdalene College, Cambridge.

Irini Katsirea, Reader in International Media Law, Department of Journalism Studies, University of Sheffield



Mathias Felipe is co-editor of the book *Journalism, Data and Technology in Latin America*, published by Palgrave Macmillan in 2021. He was previously a researcher at the University of Navarra, Spain, and a visiting researcher at the Queensland University of Technology in Brisbane. His research interests include the changing nature of communications driven by technology innovations, particularly in journalism and online social networks.

Mathias Felipe de Lima Santos, Post Doctoral Researcher, Digital Media and Society Observatory, Federal University of São Paulo



Lungile's research interests are in photography, memory, journalism and genocide studies. His work has been published in local and international journals such as Journal of Genocide Research, African Journalism Studies, Critical Arts, Nations and Nationals, Journal of Communication Inquiry, and Media, Culture and Society.

Tshuma Lungile Augustine, Post-Doctoral Research Fellow, Department of Communication and Media, University of Johannesburg



Ross writes widely on the NHS, London politics and transport. He has been a journalist for 30 years. He has also been editor of the Ham&High and political editor of The Scotsman.

Ross Lydall, Health and City Hall Editor, London Evening Standard



Susan co-chairs the Center for Data, Media & Society at Columbia's Data Science Institute. She is the author of two books: Information Security Essentials: A Guide for Reporters, Editors and Newsroom Leaders, and Practical Python: Data Wrangling and Data Quality. Her research centers on security and privacy issues affecting journalists and media organizations. Prior to Columbia University, Susan was senior programmer for the Wall Street Journal News Graphics team.

Susan E McGregor, Co-Chair, Center for Data, Media & Society, Data Science Institute, Columbia University



Claire Miller is an award-winning data journalist & expert on FOI, with over a decade's experience of using data skills to dig out data. find stories and create visualisations.

Claire Miller, Data Team Lead, Reach plc



Pete is the editorial lead with the BBC's Shared Data Unit, a collaborative journalism team that provides public interest stories to a network of local news outlets across the UK. The team also offers data journalism training opportunities throughout the year. He has worked at the BBC since 2010 in a variety of digital roles.

Peter Sherlock, Assistant Editor, BBC Shared Data Unit



Nathan is the chief executive of Hacked Off, the campaign for a free and accountable press. He is a board director of the Media Reform Coalition, which campaigns for ethical journalism, media plurality and to protect investigative and local journalism. He is also a board director of the Independent Media Association, which is an association of independent news publishers committed to high standards.

Nathan Sparkes, Chief Executive, Hacked Off



Dr. Florian Stalph's research focuses on data and computational journalism. Previously he was a research assistant at the Centre for Media and Communication (ZMK) of the University of Passau, Germany. As well as data journalism, his research focuses on related forms of digital journalism and the implications of digital media practice and technology for the journalistic field.

Florian Stalph, Post Doctoral Research Fellow, Department of Media & Communications, LMU Munich



Sina is a doctoral candidate and research assistant in the Department of Media and Communication (IfKW) at LMU Munich. Previously, she obtained a master's degree from the Institute of Media and Communication Studies at Freie Universität Berlin. Her research focuses on automated journalism as well as related forms of digital journalism.

Sina Thaesler-kordonouri, PhD Student, Department of Media & Communication, LMU Munich



Jingrong has been working in the area of journalism and news media for 20 years, first as a journalist then as an academic. She joined the Department of Journalism Studies at the University of Sheffield in September 2019. Previously, she was Senior Lecturer in Digital Media and Culture at Brunel University (2016-19) and Lecturer in Media and Communication at the University of Leicester (2009-16).

Jingrong Tong, Senior Lecturer in Digital News Cultures, Department of Journalism Studies, University of Sheffield

Vaccinating data journalism: applying lessons from the Govid-19 pandemic to a new phase in the news cycle

Louis Goddard

The Covid-19 pandemic undoubtedly boosted the profile of data journalism. As a near-simultaneous global news event with severe ramifications for both public health and the economy, it was ideally suited to data-driven analysis. The wide circulation of interactive data dashboards from public, private and academic organisations, as well as the phenomenon of 'viral' pieces of ad hoc data analysis circulated on Twitter, have been recognised as important new developments in the way data journalism is produced and consumed. But is this shift contingent on the particular characteristics of the pandemic as a news event, or can lessons learned over the past two years be used to keep data journalism front and centre as the global news cycle moves on?

This paper will review modes and forms of data journalism that have emerged, broadly speaking, over the course of the pandemic – chart-heavy Twitter threads, analytical Substack posts, data-driven supply chain monitoring, Twitter Spaces, etc. – and assess their applicability to Russia's invasion of Ukraine, drawing on the author's own experience of working on a team tasked with producing quick-hit investigative work in response to this new crisis. In particular, it will argue that data-driven reporting on the economic and financial consequences of the Covid-19 pandemic has primed audiences for news content that takes an analytical lens to the complexities of geopolitics and the global economy, but that the optimal form and venue for this content remains a subject of debate.

Sky News: Automating the COVID-19 story in the newsroom Carmen Aguilar García

The pandemic has been an important drive for data journalism in Sky News and a challenge since the first cases and deaths were reported at the beginning of 2020. The number of data-driven stories as well as the rhythm of publication rapidly increased, with many of these stories among the most read. Data journalism has become essential in Sky News and it is now at its core. But data journalists have dealt with dozens of datasets that update daily and a growing list of data sources. And they have also had to familiarise with the political and science fields relatively quickly, as it has been a highly demanding story that has evolved very fast. To face up to this challenging context and to be able to respond to the demand, Sky News developed a system to automate part of the work involved in the coronavirus story. It periodically gathers, cleans and structures data from dozens of sources, as well as performs calculations and automatically visualizes data in different platforms without the need of any

human intervention. This project has been crucial at Sky News - and especially for the data team. It has contributed to ensuring consistency across departments, making the work of data journalists more efficient and faster, and allowing for more creative content. All of this ultimately impacts on higher quality content for our audience.

Data journalism and the Covid-19 pandemic: analysis of data dissemination strategies across broadcast, online and print news media

Farzeen Heesambee

This research aims to advance existing research in digital journalism and the use of data-driven journalism during crisis reporting to demonstrate the interplay between the science community and journalism. Previous research around pandemic reporting did not focus on the role of data in reporting (Pieri, 2018; Perreault and Perreault, 2021), with COVID-19, the focus of reporting has predominantly been based on data visualisation. Throughout the pandemic, health statistics have been propelled as the driving force in journalism. Journalists have acted as 'connectors' (Pentzold et al., 2021) drawing on a range of authoritative sources such as 'healthcare professionals' to create a sense of accuracy through scientifically validated information (Perreault and Perreault, 2021). However, widely reported confusion and uncertainty surrounding some of the key UK Government's messages, as well as differential policy and strategies within local authorities, have highlighted the need to look closely at how scientific data, its mediation and public policy are negotiated and understood within the wider public sphere. Different news outlets opted for different sources of information with contradictory data, all labelled as covid data, thus making navigating through the processes of data visualisation and representation an incomprehensible disarray. The research will adopt a data life-cycle approach to COVID-19 data and integrates this with policy analysis and framing within media.

'When Two Strangers Meet': An Uptake of Data Journalism by Investigative Journalists in Zimbabwe in the Covid-19 era

Tshuma Lungile Augustine

The COVID-19 pandemic reconfigured many industries owing to the need to adjust and remain afloat during a global crisis. Thus, the outbreak of COVID-19 affected the global journalism industry in different ways: (1) there was an avalanche of data from different stakeholders, (2) whistle blowers released or leaked more information and data around the pandemic, and (3) audience were put on hard locked down, and this meant spending much of the time on the news trying to get more information about

the pandemic. Against this background, at the height of the COVID-19 pandemic in Africa, newsrooms produced investigative pieces with some of the stories being based on data that could have been supplied by whistle blowers. Examples include the Digital Vibes scandal in South Africa and the Covidgate scandal in Zimbabwe. Given these circumstances, journalists had to use the opportunity, especially the availability of data to produce their stories. Hence, this study seeks to explore how the selected media start-ups in Zimbabwe. The News Live and The Morning Star used data journalism in investigative stories. These two forms of journalism are 'strangers' in the journalism field in Zimbabwe because they are less developed and under searched and their entanglement during the pandemic calls for further exploration.

Exploring audience perceptions of, and preferences for, data journalism Neil Thurman, Florian Stalph & Sina Thaesler-kordonouri

Although data journalism has increased in volume and visibility, little is known about how it is perceived and evaluated by audiences. This study helps fill this research gap by analysing the characteristics of textual data journalism that a diverse group of 31 news consumers pay attention to and, within those characteristics, where their preferences might lie. In eight group interviews, participants read and discussed articles chosen to represent the diversity that exists in the forms and production of data-driven journalism. Our analysis reveals 28 perception criteria that we group into four major categories: antecedents of perception, emotional and cognitive impacts, article composition, and news and editorial values. Several criteria have not been used in prior research on the perception of data journalism. Our criteria have obvious application in future research on how audiences perceive different types of data journalism, including that produced with the help of automation. The criteria will be of interest too for researchers studying audience perceptions and evaluations of news in general. For journalists and others communicating with numbers, our findings indicate what audiences might want from data-driven journalism, including that it is constructive, concise, provides analysis, has a human angle, and includes visual elements.

UK audiences' perceptions of COVID-19 data visualisations: a case study

Jingrong Tong

Data visualisations have been prominent in the COVID news coverage during the COVID-19 pandemic. They are usually thought to be an effective way of presenting COVID data to audiences to engage them and help them understand relevant issues. However, is it really the case? How do audiences read, understand and trust data visualisations? What factors may influence their reading and understanding of data visualisations? To answer these questions, this paper discusses UK audiences' perceptions of COVID-19 data visualisations, drawing on the author's interviews with UK audience members from diverse backgrounds. It discusses how UK audiences perceived key aspects of selected data visualisations such as likeability, as well as the factors that influenced their perceptions. It draws our attention to the vital role played by key influential factors such as audiences' demographic backgrounds and data literacy levels in influencing their perceptions of COVID-19 data visualisations. It also explores the implications of these findings for data journalism practice and reflects on the limitations of the research.

Data Journalism and Audience Engagement in 34 Countries: Instrumental or Substantive?

Jason A. Martin, Lindita Camaj and Gerry Lanosga

Digital innovation in data journalism is driving new ways of thinking about audience, necessitating the need for more research attention to how data journalists incorporate engagement in their work and whether news organisations conceive of their data journalism audience as active users or reactive consumers (De-Lima-Santos & Mesquita, 2021; Lopezosa et al., 2021). Previous studies suggest that audience engagement practices in journalism are conditioned by a range of external factors that tend to diverge into instrumental or substantive forms of audience engagement (Lawrence et al., 2018). Yet, we need more scholarly attention to cross-national differences in data journalism practice with focus on journalistic culture, media market structures, and political context (Appelgren et al., 2019).

This study attempts to address those problems through analysis of data gathered from in-depth interviews with data journalists from 34 unique countries during the first months of the COVID-19 pandemic. Participants were asked how they define their professional roles in relation to their audience, how they value data transparency in their reporting methods, how they integrate citizens into their professional routines via crowdsourcing, and how they conceptualize audience engagement via visual interactivity in data driven stories. Results contribute to a better cross-national understanding of how data journalists think about audience engagement with attention to the contextual factors that influence their work, including public data transparency environment, organisational structure and goals, and epistemology about journalism's role in their local communities. Results echo sentiments in recent debates that emphasize the need for audience-focused news that empowers the public.

Credibility, objectivity, and transparency of data journalism during COVID-19: The audience's expectations when data leads to misleading information

Mathias Felipe de Lima Santos and Marília Gehrke

Media trust is vital for overall social cohesion, as it moderates the relationship between media users and content, thus helping to build more robust democracies (Kotta and Limayeb, 2016; Usher, 2018). Scholars have pointed out several reasons for this increasing distrust, such as fragmentation of media (Kovach and Rosenstiel, 2011) and the information disorder (Wardle and Derakhshan, 2017). Thus, publishers seek new ways to regain audiences' confidence, such as data journalism. The practice became a 'silver bullet' to expose the covert operations of the political and business

elites, promising objectivity, transparency, and trustworthiness (Anderson, 2018; Hermida and Young, 2019). However, data are human-made products, subject to error, and reflect discriminatory behavior in their production. Gitelman (2013) emphasizes there is no "raw data" and that it is necessary to carefully check how it is produced (Lesage and Hackett, 2014).

In 2021, the newspaper Folha de S. Paulo published data stories denouncing the application of expired COVID-19 vaccines. The articles indicated that 26,000 doses of AstraZeneca's immunization were out of date when applied in more than 1,500 municipalities, based on official data offered by the Ministry of Health (Gamba and Righetti, 2021). Generating a buzz on social media and in the section of comments on the newspaper's website, this became the most accessed news story in its history (Mariante, 2021). The divergent information generated much noise. Folha de S. Paulo took a week to correct the information. It indicated that the errors were caused by many flaws in the insertion of COVID-19 vaccination records and duplication of data entry, reducing journalists' errors. Thus, the following hypotheses have been evaluated: (H1) In the scenario of misleading data, the audience associates a manmade error with disinformation, and (H2) In the perception of error, the audience will claim well-practiced fundamentals of journalism.

Drawing upon data journalism, trust, and disinformation theories, our study analyzed the audience's perspective about the case of one of the most important newspapers in Brazil under the lens of journalistic values: credibility, objectivity, and transparency (Karlsson, 2020; Shapiro et al., 2006). Our study applies a mixed-method approach to evaluate these hypotheses based on users' comments on three news articles published about the case. In total, 161 comments were collected and analyzed by three coders. From our dataset, only 74 comments were considered appropriate for the methodology proposed in this study.

Results show that almost half of the comments were related to credibility (47.30%), while objectivity (32.43%) appears in second place. Surprisingly, our analysis indicated that there are claims for journalistic values when the audience is confused by the information provided. Conversely, some readers have associated the error with "fake news," which may increase the distrust of journalism and cause credibility loss. Our findings also cast light on the importance of indicating the limitations of the data, the methodological processes, and even creating repositories with the codes used so that other people can replicate and check the analyses. Our article concludes by stressing the importance of corrections in developing an effective strategy for reversing the deterioration of media trust and combating the association of journalism with disinformation.

Supporting long-haul reporting: Data integration and management in the newsroom

Susan E McGregor and Joanna Kao

Newsrooms thrive on the independence and initiative of their journalists, but these same qualities can lead to duplicated work and inconsistent coverage for long-haul, data-intense news topics like COVID. Even today's most experienced newsrooms struggle to coordinate data resources, terms and approaches across desks, leading to wasted effort and a potential loss of reader trust. We identify several factors that contribute to this situation, which we have observed across multiple newsrooms in the past 15 years. Specifically, we note that employee churn, lack of buy-in from leadership, outdated attitudes towards certain types of journalistic work and a failure to recognize and cultivate quality data as a tangible asset all contribute to fractured approaches and lost opportunity at many journalistic organizations. At the same time, we recognize that both economics and the nature of value production within a newsroom - specifically, the need to adjust priorities in real time and a cultural bias towards risk rather than investment - makes overcoming these obstacles a particular challenge. Still, we identify strategies for building and growing an integrated data practice for newsrooms that can increase journalistic quality and impact at news organizations of all sizes. We conclude by describing areas of ongoing concern and future research.

Data journalism: do we need to get back to the office to do it properly? Ross Lydall

The pandemic has changed working practices. But in journalism, it has widened the cracks between being tied to an office desk, within sight of the newsdesk, rather than created them. Many journalists believe the best stories aren't to be found in the office. Nor are they to be found at the bottom of a spreadsheet. But does data journalism benefit from being influenced by colleagues? Is it a team game? Or is data merely the blocks upon which a story is constructed or verified, but rarely the source of the inspiration for a genuine exclusive?

The BBC's Shared Data Unit: Data journalism and the public interest Peter Sherlock

The BBC Shared Data Unit has continued to promote and support data journalism at a regional level across the UK, whilst making national and local headlines with a series of public interest data investigations. The team's collaborative model sees it share its

data journalism with more than 1,000 local news outlets across television, radio, print and digital. During the course of the year, the team have also led three week-long masterclasses in investigative journalism for regional journalists and hosted more than 40 drop-in sessions for journalists wanting to learn key technical skills. The masterclasses include more than 21 hours of interactive learning.

A key part of the team's editorial remit is to find news stories that are hiding in plain sight, using data journalism techniques to explore the impact of local and national government policies on the lives of individuals. A commitment to transparency and the open data movement is at the heart of our team and our output. We have demonstrated how adopting open data principles enhances the trust in our reporting, builds a personal relationship with audiences and helps to engage them in the process of journalism. We demonstrate our commitment to transparency through sharing source data, methods and code for every project we embark upon.

Does having a TV in your bedroom lead to more sex?", and more improbable headlines backed by questionable data: The ethics of data journalism

Nathan Sparkes

Data is a critical tool in modern journalism, standing up stories or driving them directly.But journalists must be assured of the integrity of the data they publish or rely upon (or be explicit about its flaws), it should be presented accurately and transparently, and it must be set in its proper context.

Newspapers don't always use data ethically. From reporting on COVID, to climate change, to using fundamentally flawed data to stand up sensational headlines, there are dozens of examples of data being poorly used in news articles. Failure to use data ethically in journalism can have damaging consequences and is almost always a breach of the NUJ code of conduct.

By reviewing examples of good practice and bad, in different publications and from a variety of data sources, we can establish basic principles for the ethical use of data in journalism.

Data Literacy for Journalists

Jingrong Tong and Claire Miller have created a free, six-week course specifically for journalists and those who work in the news industry at all levels.

The course, available on the FutureLearn platform, will help to improve data literacy and empower users with the fundamental skills needed to understand, use and report data accurately and ethically.

Users will learn how to collect and interpret data, identify and understand the stories behind the data, and accurately share the information with readers and audiences.

Syllabus

Week 1: Data for journalists

Week 2: Collecting data

Week 3: FOI skills

Week 4: Cleaning data

Week 5: Understanding data

Week 6: Presenting data

The course requires a flexible commitment of around two to three hours per week over six weeks.

www.futurelearn.com/courses/data-literacy-for-journalists

Data Literacy for Journalists is a University of Sheffield course created in partnership with Reach plc and funded by the Higher Education Innovation Fund (HEIF) COVID Recovery Fund.

For further information, contact the course team: dataliteracyproject@sheffield.ac.uk

'Data Journalism and audience engagement in the post-pandemic era' has been organised by Dr Jingrong Tong of the University of Sheffield, in partnership with Reach plc and with funding from the Higher Education Investment Fund

Department of Journalism Studies University of Sheffield 9 Mappin Street Sheffield S1 4DT

www.sheffield.ac.uk/journalism