Storytelling Structures in Data Journalism: Introducing the Water Tower structure

Bahareh Heravi
Institute for People-Centred AI, University of Surrey
Guildford, UK
b.heravi@surrey.ac.uk

ABSTRACT
Reviewing the existing and long-established storytelling structures, this paper examines the use of the storytelling structures employed in data storytelling, specifically in the context of data journalism. For this, a large set of data stories from a variety of news outlets was collected, tagged and analysed. Accordingly, and reflecting on the results, the paper proposes a new storytelling structure for data storytelling, which addresses the unique requirements of this emerging area of study and practice, called the Water Tower structure. This proposed structure is an addition to the existing storytelling structures, and is specifically designed for and targeted at storytelling with data, with a particular focus on data journalism. While this paper is primarily focused on data storytelling in journalism, the contributions are believed to be of use and value to other domains such as Business.

Keywords: data storytelling, data journalism, storytelling with data, inverted pyramid, narrative structure

1. INTRODUCTION
Studying and analysing storytelling models and Narrative structures have been an integral part of the Arts and Humanities disciplines. When it comes to data stories, however, there exists a relatively small body of literature in this regard, mostly focused on the structure of data stories told through visual forms, i.e. visual narratives (e.g. Segel & Heer (2010) and Weber et al. (2018)). These studies analyse the stories which are entirely, and only, told through visual means and data visualisations. However, there is little study into the storytelling structures of day to day data stories, particularly when it comes to data stories as a product of data journalism.

On the other end of the spectrum, there is a body of literature that studies various characteristics of data storytelling, with a particular focus on the use of graphics and data visualisation in data journalism (e.g. (Loosen et al., 2017; Stalph, 2017; Ojo & Heravi, 2018; Weber et al., 2018; Young et al., 2018; Heravi & Lorenz, 2020; Stalph & Heravi, 2021)). In the context of data journalism, however, only a small fraction of data stories are told entirely and exclusively through visual means. In such scenarios, data visualisation is part of the structure of the story, but the overall structure of the story, or the narrative, is not told by one single large and sophisticated data visualisation. Instead, when used as a means of communication in journalism, data visualisations are often employed within text, including simple forms of data visualisations, such as bar charts, line charts and maps. This becomes even more prevalent when it comes to news organisations, with limited resources. In fact there exist data stories that employ no data visualisations, while the story itself is rooted in data.

The existing studies on visual storytelling and visual narratives could be used to inform the data storytelling space. However, further research is needed into the structure of data stories as produced in the day-to-day data journalism, from a variety of newsrooms in terms of size and data skills. This calls for a study that focuses on the structure of data stories in all forms and shapes, including those that are fully visual, but also those that have a mix of text and visuals, or even those that employ no visualisations. This paper aims to make a contribution toward addressing this gap.

1.1 Scope of this study
In this paper, I review the existing and long-established storytelling structures relevant to journalism. Building on that, I study the use of the storytelling structures employed in data storytelling, specifically in the context of data journalism. For this analysis, I examine 118 data stories published in a variety of news outlets, and provide a descriptive account of the results. Informed by the results, I propose a new storytelling structure for data storytelling, which addresses the unique requirements of this emerging area of study and practice, the Water Tower structure. This structure is an addition to the existing structures and narrative models, specifically targeted, designed for, and most useful for new models of storytelling with data, with a particular focus on data journalism.

Acknowledging various definitions at hand for terms around storytelling, the debates and the varying viewpoints, within and in between disciplines, in this paper, I use the term ‘story’ as an overarching concept, including both sequential and consequential stories (narratives), as well as “a report in a newspaper, magazine or broadcast; an item of news” (OED, n.d., p. Story: 5c), which itself may or may not be sequential. These two uses of the term story are not mutually exclusive. I further use ‘narrative’ as one way of telling a ‘story’ -- the sequential way, with a dramatic arc. This corresponds to the Narrative style of storytelling, which is further discussed in the Storytelling Structures section. Finally, the ‘plot’ is about how this sequence of cause and effects is put together and presented to the audience, often aiming for maximum impact.

2. STORYTELLING STRUCTURES
Broadly speaking, there are two oft-noted and used styles in the ways in which stories are constructed, particularly when it comes to journalism and mass communication. McKane (2006) identifies these two as *‘in medias res’* approach and the *‘chronological’* approach.

In the in medias res approach, the story starts in the middle of things, or *‘in medias res’,* at the dramatic moment of the story, and then moves on to the background and less important details. This

approach leads to what is commonly known by journalists as the *'Inverted Pyramid'* style or structure.

Chronological stories start at or before the beginning of a series of events, then go through a sequence of events in time and work their way to the end. More commonly this approach is referred to as the *'narrative'* style or Narrative structure by journalists.

2.1 The Inverted Pyramid structure

Almost all journalists are taught to structure a story in the form of an inverted triangle, or pyramid, putting the most important and most dramatic information on top, and working their way through with further details, background, information, facts and figures, etc., as they move down in the story (Figure 1).

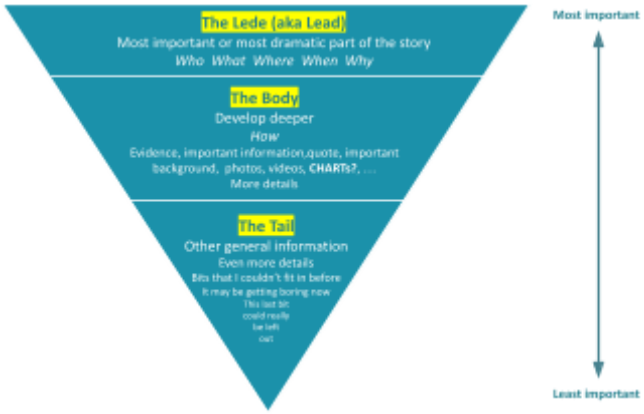

Figure 1. The Inverted Pyramid structure

The main advantage of the Inverted Pyramid structure is to grab the attention of the audience early on, and hope that would keep them engaged for the rest of the story. But if they don't remain engaged, or stop reading, they would have got the most important part of the story from the small amount of time they spent on it. Another, and perhaps the original reason for the popularity of the Inverted Pyramid approach is the way print journalism was originally produced, as it allowed the story to be cut safely from the bottom in the late hours of the night if space was an issue or to open up space for last-minute additions such breaking news stories.

While the Inverted Pyramid structure is widely used by journalists and considered an effective means of communication of journalistic stories, it has its own critics. For example, Don Fry of the Poynter Institute, who directed the Institute's writing program from 1988 to 1994, argues that the reader will no be able to fully comprehend a story written according to the Inverted Pyramid structure, as "the background goes at the bottom, somewhere between 'boring' and 'dull'. Without background, readers cannot understand the story, and simply give up before they get to the information they need" (Fry, 2004). Fry believes that "the inverted pyramid is the worst form ever invented for explaining something to another person in words" (ibid).

2.2 The Narrative style/structure

Another commonly known structure for telling stories in journalism and mass communication is the Narrative structure, which would be considered a chronological story as defined by McKane (2006).

Introduced by Aristotle, the Greek Philosopher, over two thousand years ago, the Narrative structure has its roots in ancient Greek tragedy, and has traditionally been mostly used in literary texts, but it has found its way to news and journalism. *Aristotle's Tragedy Structure* introduced a type of a plot that had a basic structure with three primary parts, or three acts: *beginning*, *middle*, and *end* (Figure 2). Aristotle's three acts are also referred to as *setup*, *conflict*, and *resolution*.

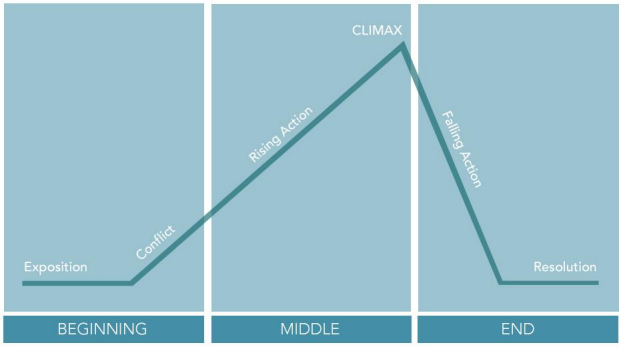

Figure 2. The 3 Act Plot's Dramatic Arc

Despite the simplicity and ancient origin of Aristotle's model, his narrative structure continues to have a significant influence on modern storytelling. This model has been used and refined over the years, and has formed the basis of other narrative structures over time.

Reflecting on the trade-off between these two widely used storytelling structures, Dyke (2020) suggests adding a data trailer to the beginning of your Narrative style story. The data trailer in fact acts the same as the Lede in the Inverted Pyramid structure, and creates what Dyke calls a Hook, which basically is a major spoiler and presents early information from the climax. It is intended to pique the interest of the audience -- or what Dyke calls an impatient executive -- and get their permission to tell them the full story (Dykes, 2020). This could provide a happy medium between these widely used traditional structures. But the Inverted Pyramid and Narrative structures are not the only two storytelling structures used in journalism.

While the Inverted Pyramid provides an excellent structure for grabbing the attention of the audience and giving them as much as you can before they lose interest or attention, the Narrative structure provides a better opportunity for deeper analyses, and continued engagement, which are often desired in investigative stories. The idea of the Lede, a hook to grab the attention of the time and attention-starved audience and bring her in, is a no-brainer in the social media age. At the same time, data journalism is all about information, exploration and explanation. While it may be useful to organise information from most important to least for time-crunched readers, a good data story, and specifically, longer data-driven investigative stories, needs a clear and logical order of cause and effects. Indeed, a mix of the two structures may be what would make a more appropriate structure for data-driven investigations.

2.3 The Martini glass

The Martini glass structure, also known as the Hourglass, is precisely one such hybrid model: it is a hybrid of Inverted Pyramid and chronological/Narrative structure. A story in this model starts with an Inverted Pyramid style summary, including the most important facts in the story. Then it switches to a chronological narrative, detailing the events step by step in a chronology. If possible, end with a kicker, a surprising twist or a

strong closing quote. Figure 3 depicts a schematic representation of this structure. This structure is, perhaps, the closest to what Dyke (2020) draws for the impatient executive by adding a data trailer to the Narrative style.

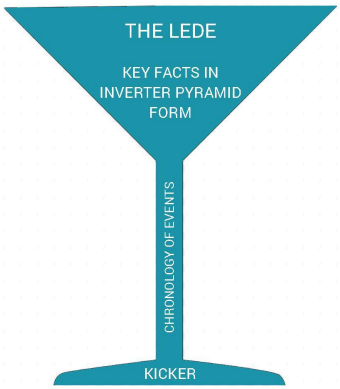

Figure 3. Martini glass structure

2.4 The Kabob

The Kabob structure, also known as the Focus style or the Wall Street Journal formula, is a structure where a story starts with a juicy Lede as an anecdotal hook, often about a person, place or event, and then progresses to a nut graf (also spelt as nut graph). Nut graf is a paragraph that provides a summary of the essence of a story without divulging every detail. It gives the reader an idea of what they can expect in the rest of the story, and why they should care about it. It also provides a transition from the lead and explains its connection to the remainder of the story. The rest of the story develops after the nut graf in as many paragraphs as needed, putting meat after meat on the Kabob skewer until it reaches the end. The story is then often concluded by returning to the person, place or event it started with, and a final anecdote (Figure 4).

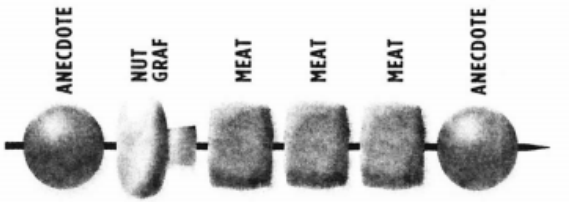

Figure 4. The Kabob structure. Image borrowed from (Harrower, 2013)

2.5 The Stack of Blocks

Discussing effective writing, Fry (2009) suggests that the following characteristics are what 2,500 years of rhetorical studies and research suggest for writing that reaches maximum power and understanding:

- A beginning, middle, and end.
- The beginning predicts the middle in form and content
- The middle contains the information, divided into sections by subjects, in logical order.
- The ending gives a sense of closure.
- The sooner readers know what the piece is about and why they should read it, the more they will understand.

Accordingly, he proposed a new structure that he calls the *Stack of Blocks*, which in his opinion caters for the above characteristics (ibid). Similar to the Narrative style, the Stack of Blocks structure has a beginning, middle, and end. Fry suggests that "the middle contains the information grouped by subject matter into *parts* arranged in a logical order (Figure 5). The beginning predicts the middle in form and content, and the ending cements the main points into the readers' memories" (Fry, 2004).

Figure 5. The Stack of Blocks structure

The main difference between the Stack of Blocks and the traditional Narrative structures is the absence of a strong dramatic arc in the Stack of Blocks structure. While the Narrative style has strong attention to the characters and their journey and the climax, the focus of the Stack of Blocks is on the development of the story through discussion of the subject matters within the story in a logical manner, as opposed to the chronology of the events. This makes the Stack of Blocks an appealing structure for stories told in investigative and data journalism.

One last note to point out here is that the Lede in the Stack of Blocks structure is not the same as Lede in the Inverted Pyramid. Rather, it provides an introduction to the story, it is the Beginning.

This section reviewed some of the most commonly used and studied storytelling structures when it comes to journalism and mass communication. The question that arises here is which of these storytelling structures are used by (data) journalists, or most suited for stories that originate from data. In other words, what are the storytelling structures when it comes to storytelling with data?

3. STORYTELLING STRUCTURES IN DATA STORIES

In this section, I examine the use of the above storytelling structures for data storytelling. To do this, a dataset of 118 data-driven stories, i.e. data journalism, published by news media organisations were curated and analysed. These include the more sophisticated data stories such as the winners and shortlisted for the 2020 SIGMA Data Journalism Award and stories from larger mainstream news outlets who produce data journalism output on a consistent basis, e.g. the Financial Times, the Economist Graphic Detail, BBC Shared Data Unit, ProPublica, New York Times, the Upshot, FiveThirtyEight, Bureau Local, the Pudding, Sky News, the Independent, Chicago Tribune, and The Washington Post. Other data stories included were from lower-tech and/or smaller or local news organisations or those that do not have data teams or partner with other news organisations for producing data stories on a less frequent basis. These include stories from Elle Magazine, New Jersey 101.5, centralmaine.com, Sunderland Echo, Liverpool Echo, Hereford Times, Berkshire Live, Naples Daily News, Times Free Press, and Alaska Public Media. These 118 stories in the dataset were tagged with one of the storytelling structures introduced earlier. The tags are Inverted pyramid, Narrative, Martini glass, Kabob, Stack of Blocks, Hybrid (of

these). If the story doesn't fit in any of these structures or a hybrid of them, it is tagged as Other.

The study reveals that amongst these storytelling structures, the Inverted Pyramid structure was the most commonly used structure for data storytelling, with 33.6% of all data stories studies being in this structure. This shows that traditional journalism training and structures are prevalent even when it comes to data-driven journalism and storytelling. Yet, while the most prevalent, only a third of the stories followed this structure.

Narrowly following the Inverted Pyramid structure, the Stack of Blocks structure was employed in just under a third (30%) of data stories studied.

Perhaps surprisingly, the data suggests that the Narrative style, and even the Martini glass structure which has the Narrative structure in its core, are in fact some of the least commonly used structures when it comes to data journalism. The Narrative structure was employed only in 6% of data stories, and the Martini glass structure was used in only 5% of data stories. This highlights that traditional Narrative structures, even when mixed with the Inverted Pyramid in the form of the Martini glass structure, are simply not used in, or not appropriate for data storytelling in journalism and mass communication.

The Kabob structure was used in just 6% of data stories, which put it at the same level as the Narrative structure. Stories with a hybrid structure formed only 3% of all data stories studied.

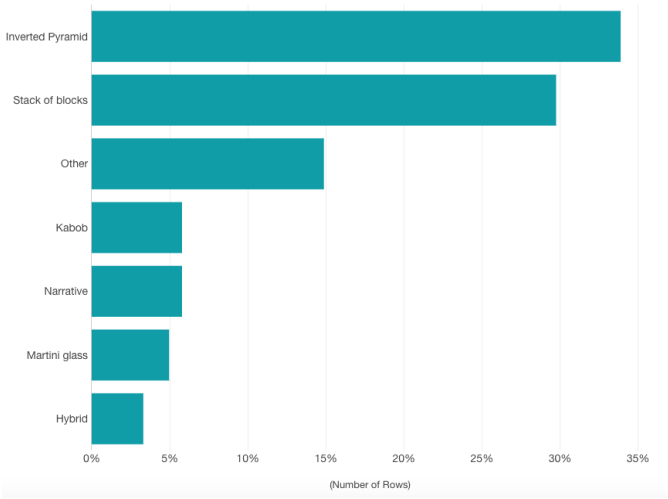

Figure 6. Storytelling structures used in data stories, *N=118*

We can observe here that while the Stack of Block structure provides a structure not too far away from the Narrative style, it is a far more used structure than the Narrative structure when it comes to data storytelling. This suggests that while telling stories with a beginning, middle and end structures is indeed common and widely used in data journalism, the traditional Narrative structure, with a dramatic arc, presenting the chronology of events, an exposition, a conflict, a big climax and a final resolution, is not the most appropriate structure when it comes to data journalism. Rather, the Stack of Blocks provides a structure that enables a logical flow and progression, focused on explaining factual information in a logical order and at the same time in a way that keeps the reader engaged and informed.

The results confirm Kosara's claim that "many data stories in journalism do not have a story arc, but rather present facts without much structure" (Kosara, 2017, p. 31). However, this study shows that data stories do indeed have structures, just not the Narrative structure with a dramatic arc. Accordingly, the story arc does not appear to be the answer when it comes to storytelling structures in data journalism.

4. THE WATER TOWER STRUCTURE: A NEW STRUCTURE FOR DATA STORYTELLING

Given the popularity of the Inverted Pyramid and the Stack of Blocks structures in data stories, and taking inspiration from the Martini glass structure, which was originally designed to cater for structures that did not entirely fall under one of the two most well-known traditional storytelling structures, here I propose a new model, which may be a suitable structure for many data stories in the context of data journalism. We call this the *Water Tower* structure.

In this proposed structure, similar to the Martini glass structure, I keep the Lede from the Inverted Pyramid structure, and then move on to the stem. Yet, given that the Narrative structure does not seem to be a popular structure when it comes to data journalism, I include a structure similar to the Stack of Blocks' in the stem.

This structure brings together the two most used structures in data journalism. It, essentially, results in a modified Stack of Blocks structure, where the Lede includes the key facts as in the Inverted Pyramid style, and then moves off to the Stack of Blocks style middle parts, and ends with a kicker.

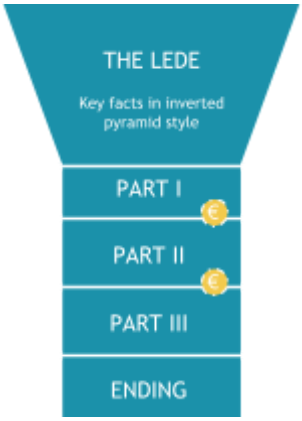

Figure 7. The Water Tower structure: a new structure for storytelling with data

The Water Tower structure starts with an Inverted Pyramid style structure, where the key information and the most important and dramatic information are presented in the beginning, as the hook. It then works its way through with further details, background, information, facts and figures, etc., as they move down to the body of the story. Unlike the Inverted Pyramid structure, however, the Water Tower structure does not keep going down to less and less important information. Instead, when it gets to the body, it exits the pyramid shape and enters a structure similar to that of Stack of Blocks, composed of a set of 'parts', where we can delve deeper into the 5Ws and the 1H and discuss different relevant topics. Utilising gold coins, we can help the reader with transitions while keeping them interested. The gold coins in data stories, and specifically in the Water Tower structure, could include data visualisation. Just like the Stack of Blocks structure, instead of leaving the background or contextual information to the end, the Water Tower structure provides the readers with the

context in little bits as needed. The ending then could be in the form of a kicker to help the reader remember. The kicker could be in the form of a conclusion, or proposing or edging towards an action, depending on the aim of the story.

The Water Tower structure could act as a suitable structure for data storytelling because it provides the flexibility to mix the two most popular structures used by data journalists, the Inverted Pyramid and the Stack of Blocks, which are the most commonly used structures by journalists for data storytelling. As such, it provides provisions for a Lede in the style of the Lede in Inverted Pyramid structure, to engage the audience and give them as much info as possible in case they do not remain engaged to the end. At the same time, it caters for a structure that allows a logical, analytical and substantive progression of the topic, enabling the journalist to delve deeper into an investigative topic, examining various influencing factors, providing evidence, and discussing any other background and relevant information.

5. CONCLUSION

The traditional storytelling structures are geared towards words, and stories produced via long-established traditional journalistic methods. The proliferation of data as a source, and the use of data visualisations as a means for communicating the final story in journalistic work, call for new models when it comes to storytelling with data. Reviewing existing and traditional storytelling structures, this paper investigates how storytelling structures are employed for data storytelling in journalism. Accordingly, it proposes a new structure specifically targeted at storytelling with data, called the Water Tower structure.

While this paper is aimed at providing data journalists with a clear and methodical view of storytelling styles for data journalism, and proposes a structure specifically for data storytelling, it does not intend to be prescriptive. The choice of the storytelling style/structure depends on the type of story under preparation, the publishing medium, and the beat, amongst other factors. For example, if a journalist is writing a short breaking news story from the latest data on housing prices, then the Inverted Pyramid may be the most appropriate structure. Similarly, if a journalist is writing about an event including individual characters, whether the character is a person, or says the life of a virus, then the Narrative structure may be your best choice. At the same time, if the story is communicating the results of a deep and long investigation on a topic, combining several datasets, going deep into analysis and investigative work, then other structures such as the Stack of Blocks, Martini glass, or the new Water Tower structure may be the more appropriate structures.

While this paper is focused on stories that are the output of data journalism and are published by news media organisations, data storytelling is not exclusive to journalism and is employed in other areas, including business. The contributions of this paper and specifically the proposed data storytelling structure could be of relevance to disciplines other than (data) journalism.

The overall storytelling structures provide valuable and practical input to storytellers and journalists when it comes to the formation of public interest stories. To add to that, there are other characteristics which could affect how a story is structured and how it unfolds and engages. Examples are the levels in which a story is author- or reader-driven, the interactivity level employed as part of storytelling, and whether it is primarily exploratory or explanatory. To provides an additional layer to how a data story is constructed, and how various characteristics could embody themselves as part of the story construction, in a follow-up study, I examine these characteristics for data stories, and investigate their inter-relationships.

6. REFERENCES

Dykes, B. (2020). *Effective data storytelling: How to drive change with data, narrative and visuals*. John Wiley and Sons, Inc.

Fry, D. (2004). *Unmuddling Middles*. Poynter. https://www.poynter.org/archive/2004/unmuddling-middles/

Fry, D. (2009, June 14). Stack of blocks. *Writing Your Way*. https://donfry.wordpress.com/2009/06/14/stack-of-blocks/

Harcup, T. (2009). *Journalism: Principles and practice* (2nd ed). Sage Publications.

Harrower, T. (2013). *Inside reporting: A practical guide to the craft of Journalism* (3rd ed). McGraw-Hill.

Heravi, B., & Lorenz, M. (2020). Data Journalism Practices Globally: Skills, Education, Opportunities, and Values. *Journalism and Media*, *1*(1), 26–40. https://doi.org/10.3390/journalmedia1010003

Kosara, R. (2017). An Argument Structure for Data Stories. *Eurographics/IEEE VGTC Conference on Visualization 2017 (EuroVis '17): Short Papers*, 31–35. https://doi.org/10.2312/EUROVISSHORT.20171129

Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2017). *Data-driven reporting: An on-going (r)evolution? An analysis of projects nominated for the Data Journalism Awards 2013 – 2016*. https://doi.org/10.1177/1464884917735691

McKane, A. (2006). *News writing*. SAGE Publications.

OED. (n.d.). OED Online. In *Oxford English Dictionary*. Oxford University Press. Retrieved 26 April 2021, from https://www-oed-com.ucd.idm.oclc.org/

Ojo, A., & Heravi, B. (2018). Patterns in Award Winning Data Storytelling: Story Types, Enabling Tools and Competences. *Digital Journalism*, *6*(6), 693–718. https://doi.org/10.1080/21670811.2017.1403291

Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 1139–1148. https://doi.org/10.1109/TVCG.2010.179

Stalph, F. (2017). Classifying Data Journalism: A content analysis of daily data-driven stories. *Journalism Practice*, *0*(0), 1–19. https://doi.org/10.1080/17512786.2017.1386583

Stalph, F., & Heravi, B. (2021). Exploring Data Visualisations: An Analytical Framework based on Dimensional Components of Data Artefacts in Journalism. *Digital Journalism*, *To appear*.

Weber, W., Engebretsen, M., & Kennedy, H. (2018). Data stories. Rethinking journalistic storytelling in the context of data journalism. *Studies in Communication Sciences*, *18*(1). https://doi.org/10.24434/j.scoms.2018.01.013

Young, M. L., Hermida, A., & Fulda, J. (2018). What Makes for Great Data Journalism? A content analysis of data journalism awards finalists 2012 – 2015. *Journalism Practice*, *12*(1), 115–135. https://doi.org/10.1080/17512786.2016.12701