Journal of Computational Social Science (2025) 8:94
https://doi.org/10.1007/s42001-025-00407-6

RESEARCH ARTICLE

Exploratory and explanatory features in data storytelling: untangling the interplay and associations with linearity, interactivity, and structure

Bahareh Heravi[1]

Received: 5 February 2025 / Accepted: 18 June 2025 / Published online: 22 August 2025

Abstract

In a journalism and communication landscape where data plays an increasingly important role in crafting reliable and engaging stories, data storytelling has become an integral part of the industry. Traditionally, journalism has often followed a linear, author-driven, and explanatory approach to storytelling. However, the rise of data storytelling has prompted a significant shift towards more interactive and reader-driven stories, with a higher potential for exploration. While the literature in the field has studied aspects such as interactivity in data visualisations, linearity and storytelling structures, the discourse surrounding exploratory features of data stories is relatively recent and scarce, and it remains largely unexplored. This paper addresses this gap and studies the exploratory and explanatory characteristics of data stories, and their interdependencies. It further examines the relationships of these facets with other key attributes of data stories, specifically linearity, interactivity, and storytelling structures. This is done through analysing a broad selection of data stories (N = 118) sourced from a variety of news outlets. Results uncover the complementary roles exploratory and explanatory characteristics play in data stories, and reveal the nuances in relationships with the other three facets. The findings provide both a deeper theoretical understanding of data story design and practical insights for data journalists seeking to balance explanation and exploration in their narratives.

Keywords Data story · Storytelling structures · Data journalism · Narrative design · Data visualization · AI and journalism

✉ Bahareh Heravi
b.heravi@surrey.ac.uk

1 Institute for People-Centred AI and Surrey Business School, University of Surrey, Guildford, UK

1 Introduction

Data storytelling has become an established practice in journalism and communication, combining the rigour of data analysis with the communicative form of storytelling. It involves both finding stories in data and telling stories with data, and is increasingly used to convey complex information in accessible ways across newsrooms and other domains.

Recent work has sought to formalise the process of data storytelling, offering workflows and design patterns to support story creation. For example, Garretón et al. [17] surveyed data storytelling approaches across academic and practitioner domains, and propose a three-stage process: *Explore*, *Make a Story*, and *Tell a Story*. They identify key elements such as interactivity, emotional effect, tool use, and narrative sequence as part of this workflow. However, they do not examine storytelling structures or characteristics in depth. Notably, Garretón et al. report that among the sources reviewed, only two referenced narrative sequence and none addressed storytelling structures directly—a gap this study aims to address.

A similar high-level workflow appears in El Outa et al. [14], who consolidate a wide range of authoring practices and proposed a conceptual model of data narratives spans from raw data to final presentation, and constitutes a four-phase model: factual, intentional, structural, and presentational. While their work provides a broad overview of data storytelling workflows, their treatment of the 'Presentation' phase—which to a large extent corresponds to what this paper refers to as *storytelling*—is limited to brief references to visualisation and interactivity. As with Garretón et al. [17], the structural composition of the story itself remains largely unexplored in their paper.

Exploring design patterns for data-driven storytelling, Bach et al. [5] propose a set of *narrative patterns* intended to guide the construction of data stories. Among these are *narrative flow* patterns, which help structure the sequencing of messages and arguments, and patterns for *framing the narrative*, which concern how facts and events are presented to support audience comprehension. In addition, they identify patterns related to *empathy and emotion* and to *engagement*, offering guidance on techniques that maintain audience interest and emotional connection. Bach et al. [5] recognise exploration as a form of active engagement and discuss how interactive elements can support it. However, their work does not offer a systematic analysis of data stories, nor does it empirically examine how specific storytelling features—such as exploratory or explanatory characteristics, or interactivity—operate within the patterns identified. A more systematic study of storytelling forms and features in data journalism could provide valuable insights for both research and practice.

Yang et al. [61] offer one of the few empirical investigations focused specifically on narrative structure in data stories. Their study specifically examines one storytelling structure, i.e. Freytag's Pyramid, which is a subset of what we consider as 'narrative structure' in this paper. They examine how Freytag's Pyramid can be adapted to the structure of data videos, proposing a design space to support narrative intent and visual storytelling. While their work demonstrates how classic narrative forms/storytelling structures can be employed in data-driven contexts, their scope is limited to a single structure. By contrast, the present study examines a broader range of storytelling structures, including narrative styles, as well as other characteristics and facets involved in making a data story, providing comparative insight into how these structures function in conjunction with exploratory, explanatory, interactive, and linear dimensions of data stories.

Together, these studies contribute valuable conceptual models and design frameworks. However, they offer limited insight into the internal composition of data stories—how different structural and functional features, such as storytelling forms, sequencing, interactivity, linearity and user agency, are deployed in practice.

Interactivity and linearity are two often studied topics in narrative studies, leading to their extensive examination within the context of data storytelling and digital journalism over the past two decades [4, 11, 47, 50, 49–54, 58]. Similarly, storytelling structures in the context of news and journalism have been widely studied in the literature [15, 19, 28, 33, 34, 36, 42, 45, 41, 64], although less so in the realm of data storytelling [25]. Two closely related characteristics to these, and emerging topics of discussion in the field, are explanatory and exploratory [8, 20, 56, 63]. Explanatory stories are those in which the author sets out the story in a specific sequence and conveys their interpretation or message in a predetermined manner [56]. They have been, to a large extent, the de facto mode of storytelling in journalism. The proliferation of data storytelling and the incorporation of data and data visualisations in journalism, however, has given rise to more exploratory stories.

Exploratory stories are those that allow users to interact with and navigate through the content in different ways, often personalising the experience or drawing their own conclusions. Exploratory stories present new opportunities for data storytelling and communication of data to the readers. While related to interactivity and linearity in story flow, the distinctive nature of exploratory characteristics in data storytelling remains under-researched. Additionally, there is a lack of research into how the exploratory and explanatory facets relate to one another, as well as to other more extensively studied features, including interactivity, linearity and storytelling structures.

Given the growing prominence of data storytelling in journalism and related fields, there is a need for empirical research that examines how stories are constructed—not only in terms of workflows and visual elements, but also in terms of the structural and functional features that shape them. Understanding the interplay between exploratory and explanatory, along with their interaction with other characteristics associated with data storytelling, such as interactivity, linearity, and storytelling structures, can help journalists better tailor their content to diverse audience needs, thereby improving engagement and comprehension.

This paper addresses this gap, through an analysis of an extensive set of data-driven stories published in a diverse range of news outlets (N = 118. It specifically focuses on two of the less studied facets: exploratory and explanatory characteristics, and further examines the relationship between these and other relevant characteristics of data stories, namely linearity, interactivity, and storytelling structures.

This study is guided by the following research questions:

RQ1: Are explanatory and exploratory characteristics positioned at opposite ends of a linear spectrum, or are they distinct characteristics that can be used in a complementary manner in data storytelling (i.e. are they independent facets that can co-occur)?

RQ2: Is there a significant association between linearity (author-driven vs reader-driven) and exploratory and/or explanatory characteristics of data stories?

RQ3: Is there a significant association between interactivity and the exploratory and/or explanatory characteristics of data stories?

RQ4: Is there an association between storytelling structure employed and the exploratory and/or explanatory characteristics of data stories?

This paper offers two key contributions. First, it develops an analytical framework for examining core features of data storytelling—including exploratory and explanatory characteristics, interactivity, linearity, and storytelling structure. The framework is then used in this study to address the research questions, and is intended to support future research and comparative analysis across contexts. Second, using the analytical framework for a systematic empirical analysis of data stories, the paper presents new empirical findings on how explanatory and exploratory features are used in data stories, and how they relate to other structural and functional aspects of storytelling. In doing so, the study applies qualitative content analysis and quantitative analysis in an integrated manner, contributing to ongoing research in computational journalism and aligning with methodological approaches in computational social science.

2 Theoretical foundations

This paper specifically focuses on five characteristics of data stories, namely interactivity, linearity, explanatory and exploratory features, and storytelling structure. This section reviews the background and theoretical foundations for each of these characteristics, which then serve as a basis for devising a codebook and analytical framework for this study.

Narratology is an established field of study that has primarily focused on analysing fictional and textual narratives, and later broadened to incorporate new ways of storytelling. In her book 'Avatars of Story', Ryan [47] extends the scope of narratology beyond traditional literary work to examine narrative structures in "new media", specifically electronic narrative forms, including the web and computer games. These forms deviate from the traditional linear narrative forms by facilitating non-linear interactive movements and explorations for readers/users [47]. Ryan [47] introduces 'interactive narratology', in which she positions interactivity as a key element of narratives for this [then] new media. Interactivity in narrative studies entails the ability of the reader/user to interact with the narrative and change its direction or course based on their preferences or standpoints, through their choice or input. Interactive features play a specifically prominent when it comes to stories involving data visualisations. Researchers including Segel and Heer [50], draw from and expand upon concepts and practices in literary narratives (e.g. [47]) to study narratives conveyed through data visualisations. As such, interactivity and linearity, being two key concepts in narrative studies, serve as fundamental features in the study of data storytelling structures and characteristics [4, 31, 35, 43, 52, 53, 58, 63].

2.1 Interactivity

Interactivity is one of the key facets of data storytelling. It transforms storytelling from a monologic process to a dialogic one, effectively turning the reader from a passive receiver to an active participant or even co-creators of content. In traditional media such as print or film, narratives are typically linear, proceeding from beginning to end in a predetermined sequence with no input from the audience. In digital and online formats, however, stories and narratives can become multi-directional, where audiences can influence and change the flow or outcome of a story, or explore multiple dimensions of it.

Ryan [47] suggests that fundamental to interactivity is the provision of 'choice' for the user. Game designer Crawford reinforces this, noting that "Every interactive application must give its user a reasonable amount of choice. No choice, no interactivity. This is not a rule of thumb, it is an absolute, un-compromising principle" (2003, p. 191). In other words, interactivity is considered the function of choice within a story, as opposed to the implementation of that granted choice.

The term 'interactivity' manifests in two interrelated ways in the literature when it concerns narrative studies and data storytelling/journalism. One usage is the broad, conceptual sense rooted in narrative and literary studies, and it refers to the overall ability of users to interact with a story. This usage of the term indeed parallels Crawford and Ryan's [47] notion of interactivity, serving as the alternative to linearity in traditional narratives. The other usage of the term stems from narrative visualisation, data visualisation and data storytelling domains, in which interactivity refers to the implementation of that choice, and is specifically entangled with interactive features/elements in data visualisations and charts. While related, it is particularly important to distinguish these in the context of this paper.

In this paper, interactivity is considered in the latter sense: it refers to the implementation of the choice through interactivity of data visualisations and other interactive elements in the story's presentation data visualisations. This is in line with Usher's definition of interactive journalism as "a visual presentation of storytelling through code for multilayered, tactile user control for the purpose of news and information" [58], p. 18). Given the focus on data storytelling and data journalism, interactive visualisations are considered to be data-driven when discussed in this paper. The more general use of interactivity—as an indicator of non-linearity in narratives through having a choice, as in Ryan's [47], is covered through the linearity facet and specifically the author-driven feature of data stories in this paper, discussed in the next section.

In summary, from this point on, the term interactivity in this paper refers to the implemented interactivity of data visualisations, unless specifically indicated otherwise.

Interactivity in data stories enables the user to interact with the data visualisation to varying degrees [3, 4, 41]. This can range from relatively simple branching scenarios where audience choices lead to different predefined routes or endings, to more complex interactive interfaces allowing users to dynamically filter data, customise their own path, or even create their own characters and storylines. Interactivity enables a more participatory, immersive, and personalised narrative engagement, allowing audiences to co-create and be part of stories, rather than just consuming them. Usher [58] highlights *tactility*, or the sense of touching, as a key feature of interactive visualisations. She suggests that tactility adds an element of play in interactive visualisations. Tidwell [57] highlights that such tactile experience can include the ability to scroll and pan, zoom, sort and arrange, search and filter and go close and far.

Examining the degrees of interactivity in data visualisations/data stories is extensively studied by other researchers and varying approaches to measuring the interactivity of data visualisations have been proposed in the literature [4, 31, 35, 43, 52, 53, 63]. Researchers in these areas have used various measures for examining the degrees of interactivity when studying data visualisations, ranging from no interactivity in static data visualisations, to limited or moderate degrees of interactivity in partially interactive visualisations or through searching, filtering and selection, to fully interactive stories and news games [43, 50, 53, 63].

Acknowledging the varying degrees and methods of measurement of interactivity in the literature, this paper focuses on the presence of interactive elements, without delving into their degree of interactivity. In other words, this paper does not study the details of interactive functionalities for individual data visualisation, such as those studied by other researchers (e.g. [35, 43, 52, 53, 63]). Rather, it examines whether or not interactivity is at all employed in the story and specifically as a particularly meaningful or significant property [24]—and not only ornamental—of data storytelling. Significant properties refer to the attributes of digital objects—in this case data visualisations—that influence their quality, usability, rendering, and behaviour [23, 24].

Taking the above into consideration, the stories in this research were tagged to indicate if the data visualisations in a given story are (a) all 'static', (b) all 'interactive', (c) a combination of 'both', or (4) employ 'animated' movement as a means of showing further details or changes in the data (but no interactivity with the data behind the animated visualisation).

2.2 Linearity: author-driven or reader-driven

Linearity in storytelling refers to the structure and control of the narrative flow. Ryan [47] discusses linearity in stories as a top-down sequential approach that guides the audience through a specified linear path from start to finish. In a linear (author-driven) story, the author controls the sequence of events and the unfolding of information. By contrast, a non-linear or multi-linear approach allows the reader/audience to interact with the story and take it in different directions based on different choices they may make, which contributes to a reader-driven experience. To study data-driven visual narratives, Segel and Heer [50] use a concept akin to Ryan's [47] interactive narratives, which they call author-driven vs reader-driven approaches

Table 1 Properties of author-driven and reader-driven stories, according to Segel and Heer [50]

Author-driven	Reader-driven
Linear ordering of scenes	No prescribed ordering
Heavy messaging	No messaging
No interactivity	Free interactivity

to narrative visualisation. They suggest that a purely author-driven approach relies heavily on the author's messaging and offers no interactivity (in neither concept and implementation senses), whereas a purely reader-driven approach "has no prescribed ordering of images, no messaging, and a high degree of interactivity" ([50], p. 1146). Their view on this is summarised in Table 1.

The characterisation of interactive and linear narratives presented by Ryan [47] emerges from text-based narratives. Segel and Heer's [50] classification, despite having been exclusively used for examining narratives conveyed through data visualisations, is itself rooted in the more traditional text-based practices of narrative studies, including Ryan's [47].

In practice, data storytelling typically resides between these two extremes, incorporating both textual and visual elements into the narrative. Thus, I argue that Segel and Heer's [50] classification is applicable for data stories that combine various means of communication, including text, visualisations, and/or video.[1] Consequently, to study the linearity of data stories in data journalism, this paper adopts the author-driven and reader-driven continuum from Segel and Heer [50] with one caveat: the provision of the author- versus reade-driven characteristics encompasses interactivity in the sense of Ryan's [47], i.e. the provision of choice in narratives regardless of the implementation. Interactivity itself is examined as its own unique facet, specific to the implementation of interactivity in data visualisationsin RQ3, as described above.

In conclusion, it is considered that data stories lie along a spectrum between the two extremes of fully author-driven and fully reader-driven. To cater for that, in this research the stories are tagged on a spectrum from 1 to 5 (1 being fully author-driven and 5 being fully reader-driven).

2.3 Explanatory or Exploratory characteristics

Explanatory and exploratory refer to the story's mode of communicating information and the role of the reader in deriving insights.

Exploratory stories offer readers the flexibility to navigate a data story in a non-sequential manner, allowing them to uncover stories that resonate with their individual perspectives and personalise their journey through various available routes. These characteristics can offer readers a sense of agency, where they can affect the

[1] Ojo and Heravi [43] use the same classification as Segel and Heer for the author-driven / reader-driven facet in the context of data storytelling, but they refer to it as the Narrative Style.

structuring of the presentation and derive their own interpretation from data [40, 56]. By incorporating exploratory elements, authors can craft data stories that adapt to readers' unique knowledge, interests, and experiences.

Exploratory stories are considered to play an increasingly important role in user engagement and their interest in stories [54, 58, 60, 63]. The freedom to explore a data story, however, requires readers to have knowledge of the topic, as well as basic literacy in data and visualisations. It also requires readers' time, attention, and cognitive investment [56].

Explanatory stories, on the other hand, often have a finite narrative structure and a prescriptive and sequential progression, with a clear beginning, middle and end [56]. In explanatory data stories, the author provides the readers with their own interpretation, context, or additional information about the story and the data at hand. Text is often the predominant method of communication in explanatory stories. When visualisations are employed in these stories, visual annotations are ordinarily used for emphasis to guide the reader's focus on areas they perceive as important or relevant to the story.

Explanatory stories can be particularly effective in engaging audiences who lack prior knowledge or interest in a given subject. These stories guide readers through complex issues or data, directing their attention to certain key points. Nonetheless, their prescriptive nature provides limited scope for readers' personal interpretation and deeper engagement. This could potentially lead to one-sided stories or those with appeal to a narrow readership.

Reflecting on the explanatory and exploratory facets of data visualisations, Young, Hermida and Fulda [63] suggest that journalism has traditionally preferred the explanatory mode when it comes to data visualisations, in the form of presenting information to the audience. However, they argue that emergent technologies have turned readers into analysts and provide them with opportunities to explore the data behind visualisation. Other research suggests that integrating explanatory and exploratory elements forms one of the most engaging approaches to storytelling [56, 60]. Thudth et al. argue that "exploratory facets are useful for data analysis and personalized navigation. Explanatory facets, on the other hand, are essential for communicating information and [the authors'] viewpoint about the data and establishing emphasis and storyline for the reader" ([56], p. 60).

Similar to author-driven and reader-driven characteristics, the explanatory and exploratory aspects of data stories have been traditionally put on the two opposite ends of one spectrum, e.g. in Barlow [6] and Weber et al. [60]. In this designation, the increase in how explanatory a story is necessarily reduces the potential for exploration of that story.

Adopting this linear perspective, Anderson and Borger-Rey [3] put explanatory and exploratory features on a linear continuum, and as direct attributes of reader-driven and author-driven characteristics respectively (Fig. 1).

Thudt et al. [56], however, suggest that data stories create new opportunities by combining these two facets. They argue that balancing exploration in a two-dimensional space, instead of the two opposing ends of a spectrum, can provide an optimum trade-off between the highly exploratory but complex and the highly explanatory but prescriptive stories, and has the potential to achieve better

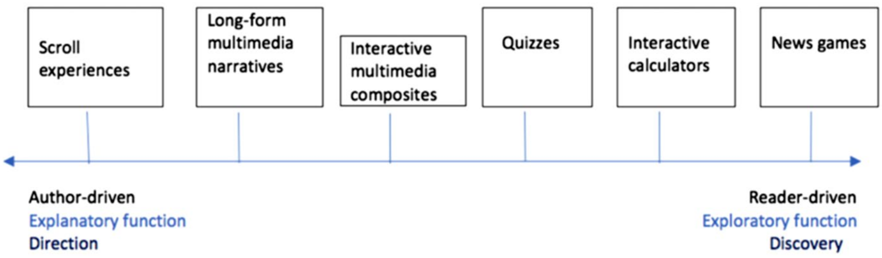


Fig. 1 From [3]

engagement and experiences for the reader [56]. Accordingly, instead of a linear spectrum with two facets at the two ends, they propose a two-dimensional space for exploration and explanation facets when it comes to data storytelling, where the two can co-exist and complement each other,see Fig. 2. Under this proposition, each data story could present a varying degree of both exploration and explanation.

In line with Thudt et al. [56], this paper takes the view that explanatory and exploratory characteristics are complementary, and have the potential of being seamlessly integrated for effective data storytelling. Yet, this perspective is not presupposed as a definitive conclusion,instead, it serves as a hypothesis subject to examination through the lens of collected and tagged data in this study, under RQ1:

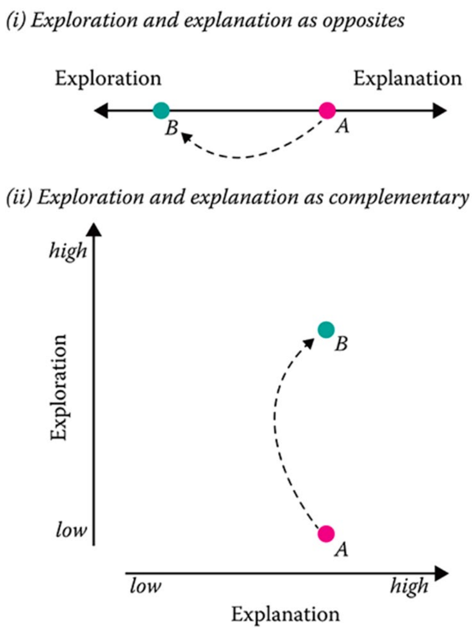


Fig. 2 Thudt et al.'s [56] position in considering exploration and explanation as complementary (ii) as opposed to exploration and explanation as opposites (i) Image borrowed from Thudt et al. [56] for presentation purposes

Hypothesis 1: exploratory and explanatory facets are complementary.

To examine if exploratory and explanatory facets are complementary, i.e. if a story or visualisation could be made exploratory while maintaining degrees of explanatory, in this paper I treat these facets as two distinct characteristics and code each story for their degree of *explanatory* as well as its degree of *exploratory* separate from one another. This enables not only the association of the two facets to be examined but also to examine the validity of Thudt et al.' [56] view that these two facets are in a two-dimensional space. If a near-perfect negative correlation between the two facets is observed, this would suggest that they are on the two ends of a linear continuum, otherwise, they belong to a two-dimensional space, and can be complementary.

Considering that exploratory stories allow readers to shape the narrative and derive their own interpretations from data, and their adaptability to the readers' unique knowledge and interests, we can anticipate that exploratory stories would facilitate a more reader-driven experience. On the other hand, with their pre-defined structure set by the author, we can expect that explanatory stories foster a more author-driven experience. This expectation, paired with the earlier assertion about exploratory features and their reader-driven tendencies, forms the basis of hypotheses that are examined under RQ2.

Hypothesis 2–1: highly exploratory stories are likely to be more reader-driven.

Hypothesis 2–2: highly explanatory data stories tend to be more author-driven.

Interactivity is an oft-discussed feature when it comes to exploratory data stories. As discussed earlier, interactivity in data visualisations allows the reader to delve into various hidden aspects of the data, and hence facilitate exploration.

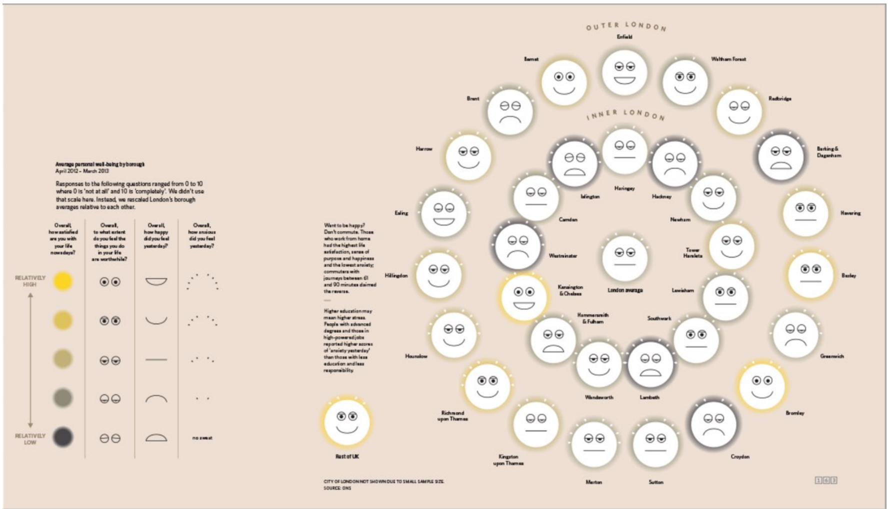


Fig. 3 An example of a static yet exploratory data-driven visualisation. From London: The Information Capital [9], pp. 162–163)

Consequently, many exploratory data stories provide interactive capabilities, particularly when dealing with larger sets of data. However, interactivity is not a prerequisite for exploration. Even static representations, such as the visualisation of the average well-being in London boroughs [9] presented in Fig. 3, can provide a moderate to high degree of exploratory characteristics, allowing viewers to ask various questions and formulate individual interpretations while providing a minimal explanation. There are many other examples of static data/information visualisation in print books, such as London the Information Capital ([9], pp. 162–163), Information is Beautiful [37], Knowledge is Beautiful [38] and Facts are Sacred [46] that provide varying degrees of exploratory features in their visualisations. Despite being static, these data visualisations allow the reader to explore the possibilities, find the story more relevant to them and draw individual interpretations, all while in some cases offering minimum or no explanation or textual narrative structure.

To examine this further, this paper studies the relationship between interactivity and the explanatory and exploratory features as part of RQ3, under which we examine two specific hypotheses:

Hypothesis 3–1: Data stories with a higher degree of exploration potential are more likely to employ interactive data visualisations, whereas those with lower degrees of exploration potential tend to use static elements.

Hypothesis 3–2: Data stories with lower degrees of explanation are more likely to incorporate interactive elements, whereas data stories with a higher degree of explanation predominantly use static elements.

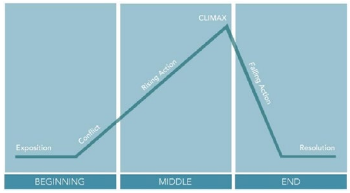
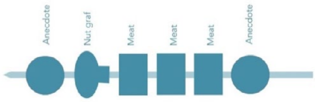
One more aspect is studied in relation to exploratory and explanatory features of data stories in this paper, and that is the storytelling structures employed, described in the next section.

2.4 Storytelling structures

Storytelling structures are a well-studied and established field of study when it comes to journalism. In the realm of data-driven stories, existing research has predominantly examined the narrative/storytelling structures or patterns solely within the data visualisations [4, 31, 35, 43, 52, 53, 63]. This prevailing trend often overlooks the broader composition of the story, which might encompass data visualisations, text, and other elements.

Research specifically addressing storytelling structures in data journalism, encompassing both text and data visualisations, remains scarce and relatively recent. Heravi [27, 25] has explored this area by examining the use of these structures in data storytelling. As part of this, she provides a set of definitions of storytelling structures pertinent to data storytelling in journalism. This paper adopts Heravi's [27, 25] definitions and classifications and therefore, refrains from providing an extensive overview of the storytelling structures. Instead, only a brief description of each structure is provided to enable comprehension of the underlying concepts needed for this paper. The reader is referred to Heravi [27, 25] for a more comprehensive discussion on each storytelling structure. Table 2 provides a schematic summary of these storytelling structures.

Table 2 A schematic summary of storytelling structures, borrowed from [22, 27, 25]

<table>
<tr><td>Inverted Pyramid</td><td colspan="2">Narrative structure</td></tr>
<tr><td>
</td><td colspan="2">
</td></tr>
<tr><td>Hourglass</td><td>Kabob</td><td>Stack of Block</td></tr>
<tr><td>
</td><td>
</td><td>
</td></tr>
</table>

The Inverted Pyramid is a commonly used storytelling structure, where, where the most important or dramatic information is presented at the beginning of the story (the lede), followed by additional details, background, information, facts and figures as the story progresses. Despite being widely used by journalists and considered an effective means of communication, enabling a deep understanding or full comprehension of the subject [16].

The Narrative structure is focused on the chronological progression of the story. A story following this structure typically has a clear beginning, middle, and end, often reaching its peak with a significant revelation that serves as the climax.

The Hourglass structure, also known as the **Martini Glass structure**, is a hybrid of the inverted pyramid and narrative structures. A story in this model starts with an inverted pyramid style summary, including the most important facts in the story, and then switches to a chronological narrative, detailing the events step by step in a chronology.

The Kabob structure begins with an anecdotal hook as a lede, followed by a nut graf, which summarises the story's essence without revealing all details, while linking the lede to the rest of the story. The story then unfolds further, adding more details, akin to meat on a skewer, until it concludes by reflecting on its beginning, often through a final anecdote.

The Stack of Blocks structure contains a beginning, middle, and end—much like the narrative structure, but the middle segment logically groups information by subject matter [16]. Unlike traditional Narrative structures, which emphasise character development and dramatic climaxes, the Stack of Blocks structure focuses on logically discussing subject matters. 'old coins' are dispersed as interesting elements throughout to maintain reader engagement.

To examine exploratory and explanatory features of data stories concerning their storytelling structures (RQ4), each data story in this paper is tagged as one of *Inverted Pyramid*, *Narrative*, *Hourglass*, *Kabob, Stack of blocks*, *Hybrid* (for stories that combine elements of the above structures in ways that don't cleanly fit one category) or *Other* when a data story does not fall under any of these structures.

2.5 Analytical framework

Drawing on the theoretical foundations above, this paper devises an *analytical framework* for examining the above facets of data stories, and addressing the research questions. Table 3 provides a summary of this analytical framework.

Table 3 The analytical framework used for coding data stories

Characteristic	Codes
Interactivity Adopted from [49, 52, 53]	N/A Static: Involving only static visuals Interactive: Involving only interactive visuals Both: Involving a mix of both static and interactive visuals Animated: Involving animated graphics
Author-driven to Reader-driven (Linearity) Adopted from [50]	1: Highly/fully author driven 2: Mainly author driven 3: Somewhere in between 4: Mainly reader driven 5: Highly/fully reader driven
Explanatory Adopted from [56]	0: Not at all 1: Very low 2: Low 3: Somewhere in the middle 4: High 5: Very high
Exploratory Adopted from [56]	0: Not at all 1: Very low 2: Low 3: Somewhere in the middle 4: High 5: Very high
Storytelling structure Adopted from [27, 25]	Inverted Pyramid Narrative Hourglass Kabob Stack of blocks Hybrid Other

3 Method

This paper examines the extent to which exploratory and explanatory characteristics are employed in data storytelling in journalism, and whether or not the two facets are complementary (RQ1). In addition, it explores how these two facets are associated with linearity (RQ2) and interactivity (RQ3) characteristics. Finally, it examines exploratory and explanatory features of data stories with respect to their storytelling structures (RQ4).

In doing so, this study uses a mixed-methods content analysis to examine the structural and functional features of data stories. It is guided by an analytical framework developed for this research (Table 3), which incorporates key dimensions such as explanatory and exploratory characteristics, interactivity, linearity, and storytelling structure. The framework was used to code a purposive sample of 118 data stories published by a range of news organisations. The resulting coded data were then analysed using descriptive and inferential statistical methods to identify patterns and relationships among features.

3.1 Data collection

To address the research questions, a dataset of data-driven stories published by news media organisations was compiled. This dataset was coded and analysed in accordance with the analytical framework discussed in the previous section, summarised in Table 3.

The starting point for the selection of data stories to be included was the SIGMA Data Journalism Awards 2020. The SIGMA Awards is the premier annual data journalism competition, celebrating the best data journalism around the world. They published the list of all submitted, shortlisted and winning entries. From this, a set of winning and shortlisted stories were included in the dataset. Given the linguistic limitations, this study focuses on articles originally written in English, or those easily translatable into English via the Google Translate plugin, as the analysis heavily involves understanding the text and content. A practical approach was adopted in this regard: if the team felt confident in comprehending the text and context using the basic translation provided by the plugin, it was retained; otherwise, it was discarded. This resulted in the inclusion of only four non-English stories/projects, three in Spanish and one in Ukrainian.

Given the high prestige and prominence of the SIGMA Awards internationally, the shortlisted and winning stories typically represent some of the most sophisticated data stories. This paper aims to examine a broad spectrum of data stories, not limited to the top-tier, award-winning pieces. To this end, a series of stories from news organisations with a consistent track record of publishing data stories and practising data journalism—e.g. those with dedicated data or graphics units or with a specific focus on data journalism in their mission statements or activities—were added to the dataset. These included data stories from the Financial Times Data, Economist Graphic Detail, BBC Shared Data Unit, The Guardian Data, ProPublica, New York Times, The Upshot, FiveThirtyEight, Bureau Local, The Pudding, Sky News, The Independent, Chicago Tribune, and Washington Post.

While the resulting list comprised a higher number of day-to-day data stories, the majority still originated from the more prominent and well-resourced organisations with more advanced technological capabilities or greater resources. To ensure the results are an appropriate representation of the industry as a whole, it is paramount to include stories from the lesser-known, technologically less-equipped teams, smaller organisations or those with limited resources in data journalism. However, locating such entities within the vast journalistic landscape poses a challenge. An approach was devised to address this by identifying established entities known for their collaborative practices and tracing their co-publication and content partnerships with smaller organisations. Notably, the BBC Shared Data Unit, ProPublica, and the Bureau Local consistently collaborate with smaller local media entities for the production and publication of data-driven stories, including co-publications and re-publications. By tracing their partnerships, an additional set of data stories was identified from smaller or local news organisations. These include 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. The final dataset included a total of 118 data-driven stories from a diverse range of 51 news organisations. A breakdown of these news organisations can be found in Appendix 1.

The requirement for understanding the content means that the final dataset predominantly comprises stories from English-speaking countries, and particularly from the US and UK. Specifically, 105 out of 118 stories originated from these two countries. Therefore, the results cannot be generalised to all languages and countries.

A breakdown of the countries of the publications involved is presented in Table 4. Apart from four stories from Mexico, Peru and Ukraine, all other stories included in the dataset were in English.

The majority of the stories included were published between 2018 and 2020, while a small number (7) were published prior to 2018. No story after 10th March 2020 was included. The reason for this was that a sudden vast number of stories including data elements and charts were being published from mid/late March 2020 because of the Corona Virus pandemic. The stories included very similar data, context and presentation styles, and impacted storytelling on any other topic. As such,

Table 4 Breakdown of the country of publication of data stories

Country (publisher)	Count	Country (publisher)	Count	Country (publisher)	Count
US	54	Bangladesh	1	Peru	1
UK	51	Canada	1	Philippines	1
Global	2	China	1	Ukraine	1
Mexico	2	France	1		
Australia	1	Greece	1		
Grand Total 118					

the decision was made to stop the data collection to a date before the pandemic. The topics covered included a very wide range including politics, social issues, economy, health, literature, music and environment.

3.2 Coding procedure

The data for this study was carefully coded by two independent coders, both with journalism backgrounds and familiarity with data-driven storytelling and data journalism. They were provided with the same codebook and underwent training to ensure consistency in their approach.

Before the full coding began, the coders undertook two rounds of pilot coding, each involving a small set of five. These served as calibration exercises to harmonise understanding and interpretation of the coding scheme, identify potential ambiguities, and ensure consistent application across cases. After each pilot round, coders compared their results and resolved discrepancies through discussion, ensuring a shared coding login and consistent and reliable coding process. Once calibration was completed, the coders proceeded to code the full dataset independently.

To assess the inter-coder reliability, Cohen's Kappa (for categorical variables) and Weighted Kappa (for ordinal scales) were calculated for the variables coded by the coders. The results demonstrated high reliability across all variables, indicating strong agreement between the coders. Specifically, the weighted Kappa values for 'Explanatory', 'Exploratory', and 'Story Type' were 0.962, 0.924, and 0.956, respectively. For the categorical variables, the Cohen's Kappa values for 'Storytelling Structure' and 'Interactivity' were 0.934 and 0.892, respectively. These scores far exceed commonly accepted thresholds, underscoring the consistency and reliability of the coding. These values underscore the consistency and reliability of the coding process, providing a solid foundation for further analysis. This is presented in Table 5 below.

After coding, any differences were resolved through discussion and consensus. In cases where consensus could not be readily achieved, a third reviewer with oversight of the research acted as an adjudicator and made the final decision. The final dataset used for analysis thus reflects these agreed-upon codes, ensuring both interpretive depth and methodological rigour.

The use of two independent coders is a widely accepted and established practice in qualitative and mixed-methods research, especially when supported by a clear and

Table 5 Summary of intercoder reliability for examined variables

Variable	Test	Intercoder reliability
Explanatory	Weighted kappa	0.962
Exploratory	Weighted kappa	0.924
Story Type	Weighted kappa	0.956
Storytelling Structure	Cohen's kappa	0.934
Interactivity	Cohen's kappa	0.892

well-developed codebook, coder training, a systematic process for resolving disagreements and discrepancies, and robust inter-coder reliability measures [32, 39]. This approach, i.e. two independent coders, is also common and routinely employed within journalism, communication, and digital media research, either coding in parallel with inter-coder agreement measured (e.g., [1, 55]), or using iterative consensus (e.g., Schröder et al. [48]). Similar two-coder designs have also been used in large-scale content analyses with quantitative goals (e.g., [18, 29]), underscoring the broad acceptance of this approach across journalistic and communication research.

This approach ensured a consistent and rigorous application of the coding scheme while allowing for interpretive nuance aligned with the study's qualitative elements.

4 Results

Having established the coding framework and dataset, this section presents the findings of this study, organised by topics studied under research questions.

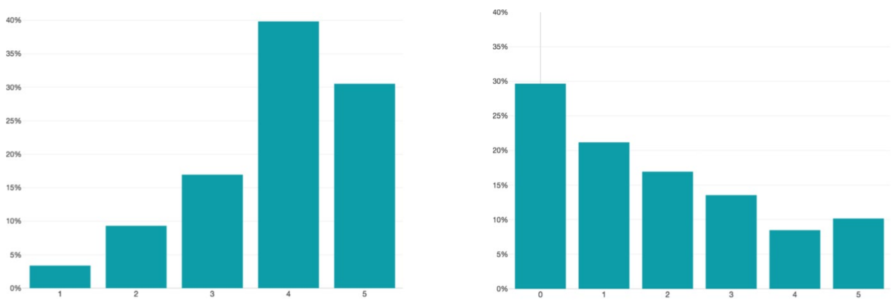
4.1 Explanatory and exploratory characteristics (RQ1)

This section commences the discussion of the findings by addressing the initial research question (RQ1) on the extent of prevalence or usage of either or both exploratory and explanatory characteristics in data storytelling. In addition, it examines whether or not these two characteristics exhibit a complementary, two-dimensional relationship, rather than existing merely as polar opposites on a linear spectrum. Understanding this interplay could shed light on how effectively data stories convey information to their audience by combining these characteristics.

As a reminder, adopting Thudt et al.'s [56] point of view which posits exploratory and explanatory characteristics as complementary, the analytical framework in this paper employed two distinct scales, ranging from 0 to 5, for each of the explanatory and exploratory facets: zero (0) denotes a complete absence of the feature, while 5 indicates a very high presence (0: not at all, 1: very low, 2: low, 3: somewhere in the middle, 4: High, 5: very high). The data stories studied were coded and analysed accordingly for explanatory and exploratory features independently, followed by an examination of the correlation between the two facets.

The results suggest that the majority of data stories in the news media are highly explanatory, and offer limited opportunities for exploration. To put it in numbers, 70% of data stories studied were highly or very highly 'explanatory' (Fig. 4-i). In terms of the 'exploratory' characteristics, 30% of stories presented no exploratory characteristics at all, and another 28% presented very low, or low degrees of exploratory features. In other words, nearly 60% of data stories had no or limited exploratory features, and just under 20% of stories had high or very high degrees of exploratory features (Fig. 4-ii).

This may suggest that similar to data visualisations, as argued by Young et al. [63], data stories overall more commonly prefer the use of explanatory approaches over exploratory ones.



i) Explanatory (0: not at all … 5: very high) ii) Exploratory (0: not at all … 5: very high)

Fig. 4 Explanatory and exploratory characteristics in data stories (N = 118)

A visual inspection of the two graphs in Fig. 4 suggests a negative association between the two facets, which brings us to the second part of the RQ1 on whether or not these two facets are at the two ends of a linear continuum, or as hypothesised, they can belong to a complementary two-dimensional space (Hypothesis 1).

To test if there is a statistically significant association between the two facets, a correlation analysis was conducted. A perfect negative correlation would indicate a linear continuum with the facets at opposing ends. Conversely, any other result would suggest that these facets co-exist within a two-dimensional space, not on a linear spectrum.

The correlation analysis between exploratory and explanatory facets revealed only a medium-weak correlation ($\rho = -0.34$, Spearman). This implies that although there is a negative association between the two facets, they do not exist at opposite ends of a single spectrum. Instead, positioning them on two axes within a two-dimensional space, where they have the potential to co-exist and complement each other, offers a more appropriate representation. This validates Hypothesis 1, and not only aligns with the viewpoint of Thudt et al. [56] but also serves as further empirical substantiation of their proposition.

4.2 Exploratory/explanatory and linearity (RQ2)

Weber ([60], p. 195) notes that "experts in the field of data visualization and data journalism often refer to the terms reader-driven… and author-driven as exploratory and explanatory". Similarly, Anderson and Borger-Rey [3] put explanatory and exploratory features as direct attributes of reader-driven and author-driven characteristics on a linear continuum. However, in accordance with Thudt et al. [56] and supported by our above analysis, I do not automatically consider author-driven and reader-driven to be synonymous, identical or consequential traits with exploratory or explanatory, despite the potential associations and similarities between them. Instead, this section examines whether there is an association between the explanatory and/or exploratory facets of data stories and their degree of linearity, ranging from author-driven to reader-driven (RQ2). Earlier in this paper, I posited that exploratory data stories lean towards the reader-driven end of the linearity spectrum

(Hypothesis 2-1), and explanatory stories towards the author-driven end side of the spectrum (Hypothesis 2-2).

Considering that we have positioned explanatory and exploratory characteristics on separate spectrums in a two-dimensional space, I will not draw a direct comparison between author-/reader-driven and exploratory/explanatory. Instead, here I explore if the linearity facet is associated with each of the exploratory and explanatory facets separately. In other words, I examine if exploratory stories present are likely to be more author-driven (Hypothesis 2-1) and if explanatory stories are also more author-driven (Hypothesis 2-2).

From the graph in Fig. 5, we can observe a visual positive association between exploratory stories and linearity. This association is further confirmed as a strong one by correlation analysis ($\rho = 0.73$ Spearman). In essence, there is a significant relationship between exploratory characteristics and linearity: the more explorable data stories are, the more reader-driven they are, and the less explorable they are the more author-driven.

When examining the relationship between explanatory and linearity facets, the analysis provides a more nuanced view. An initial visual inspection of Fig. 6 reveals that stories with a higher degree of explanation are also more author-driven. This relationship is not observed for stories with a lower degree of explanatory features. This is reflected in the correlation analysis, which only confirms a weak association between linearity and the explanatory characteristics ($\rho = -0.28$ Spearman). This weak negative correlation does not support an overall correlation between linearity and explanatory features. Nonetheless, the analysis does suggest that highly explanatory stories foster a more author-driven experience. We cannot, however, make such a claim about reader-driven stories not being explanatory. Caution must be exercised in making claims about the lack of explanatory features in author-driven stories.

In short, the analysis supports our hypotheses that highly exploratory data stories often fall on the reader-driven side of the linearity spectrum (Hypothesis 2-1), and

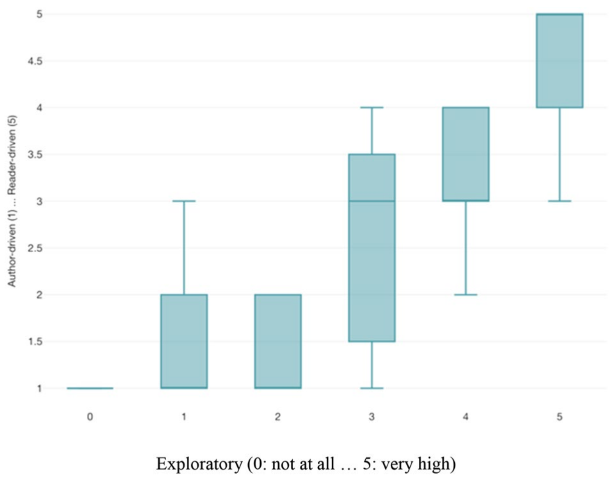


Fig. 5 Exploratory and Linearity (Correlation $\rho = 0.73$ Spearman)

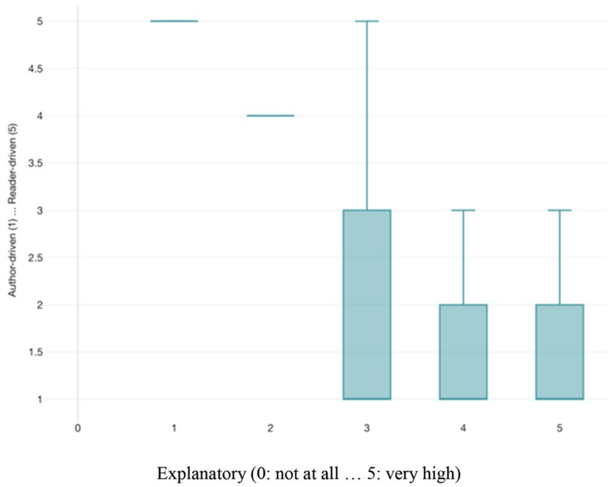


Fig. 6 Explanatory and Linearity (Correlation ρ = − 0.28 Spearman)

highly explanatory stories on the author-driven end (Hypothesis 2-2). We can also confidently conclude that stories with lower exploratory potentials follow a more author-driven approach. However, we cannot assert that stories with low explanatory degrees cater for a reader-driven experience. The fact that only a small number of data stories had low explanatory degrees, could be a reason for this low correlation.

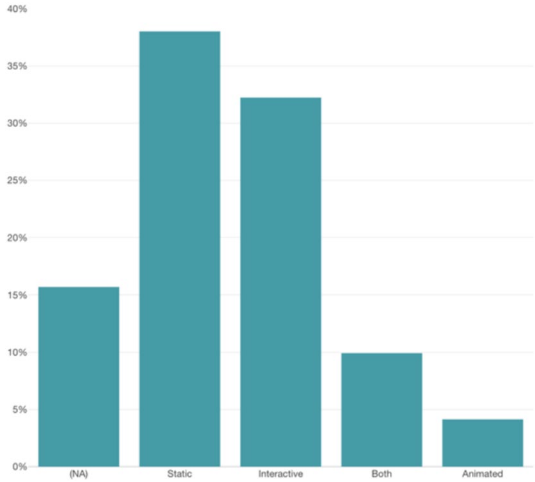
4.3 Exploratory/explanatory and interactivity (RQ3)

Exploratory stories often incorporate interactivity to allow for users' interaction as a means of exploration. Although interactivity frequently features in exploratory data stories, as noted earlier in the paper, it is not a prerequisite for exploration. This section examines the association between interactivity and the explanatory/exploratory characteristics of data stories (RQ3).

Following the analysis, it becomes apparent that a mix of both interactive and static visualisations is used in data storytelling, with neither one significantly outnumbering the other. Just over 40% of data stories presented some degree of interactivity, and just under 40% were using only static data visualisations. It is interesting to note here that a considerable number of data stories (just over 15%) did not use any visual elements in the story at all (tagged as N/A). Slightly less than 5% of the data stories were tagged as animated, denoting the use of animations—including gif animations—to show changes in graphics. These are not static objects such as images, but also do not offer options for interaction with data points or chart elements. See Fig. 7.

Having established the prevalence of interactivity in data stories, we can proceed to examine whether or not interactivity is associated with the exploratory and/or explanatory characteristics of data stories (RQ3).

Fig. 7 Interactivity in data stories (N = 118)



Focusing first on the exploratory facet, the visual inspection of the graphs in Figs. 8 and suggests that data stories with a lower degree of exploratory characteristics tend to use static data visualisations. In contrast, those with a high degree of exploration potential employ interactive data visualisations entirely.

The Kruskal–Wallis test was conducted to examine the relationship (difference between groups) between exploratory and interactivity facets (Hypothesis 3-1). The results revealed a significant difference in these variables across the groups ($H = 62$, $p < 0.001$, Kruskal–Wallis test). The p value obtained was 2.56e−12, indicating strong evidence of a relationship between exploratory and interactivity variables. In other words, the results support Hypothesis 3-1 which indicates exploratory characteristics are significantly associated with the degree of interactivity in data storytelling.

Similar to the above analysis, the Kruskal–Wallis test was implemented to assess the association between explanatory and interactivity variables. The findings

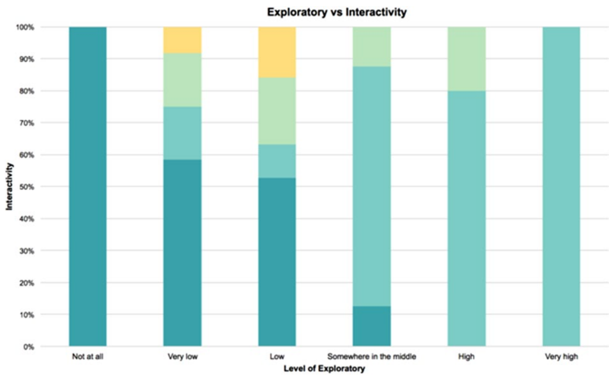


Fig. 8 Exploratory and Interactivity ($H = 62$, p value = 2.56e-12, Kruskal–Wallis test)

indicated a significant difference in these variables among the groups (H = 12, $p < 0.01$, Kruskal–Wallis test), hence supporting Hypothesis 3-2, which posits a significant relationship between explanatory and interactivity factors.

Nonetheless, further visual inspection of the data suggests that while nearly all stories with low explanatory degrees incorporate interactive elements, stories with a high degree of explanatory characteristics are not limited to the use of static elements, e.g. text and static charts. In other words, while interactive features are predominantly used for explorations, interactivity is not specific to exploratory data stories; indeed even those with a high degree of explanation, occasionally employ both static and interactive visualisations (Fig. 9).

4.4 Explanatory/exploratory and storytelling structure (RQ4)

This last section of the Results section examines the relationship between storytelling structures used in data storytelling and the employment of exploratory and explanatory features (RQ4).

The study shows that amongst the storytelling structures introduced earlier, the Inverted Pyramid structure is the most commonly used structure for data storytelling (33.6%), narrowly followed by the Stack of Blocks structure (30%). Figure 10 depicts the distribution of storytelling structures used in data stories studies.

Examining RQ4, the Kruskal–Wallis test was conducted to investigate the relationship between story structure and both explanatory and exploratory facets.

In terms of the explanatory facet, the results demonstrate a significant difference across different story structures (H = 29, $p < 0.001$, Kruskal–Wallis test). A p value of 6.66e-05, confirms strong evidence of a relationship—in terms of difference between groups—between story structure and explanatory traits. Figure 11 provides a view of the relationship between story structure and the degrees to which they are explanatory.

The Narrative style stands out amongst other structures as the only structure that entirely results in highly explanatory stories. This is perhaps not surprising as the reader is being taken through a predefined storyline in the narrative structure, with

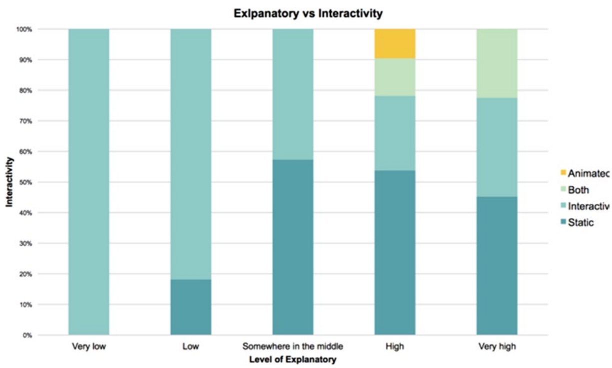


Fig. 9 Explanatory and Interactivity (H = 12, p value = 0.007, Kruskal–Wallis test)

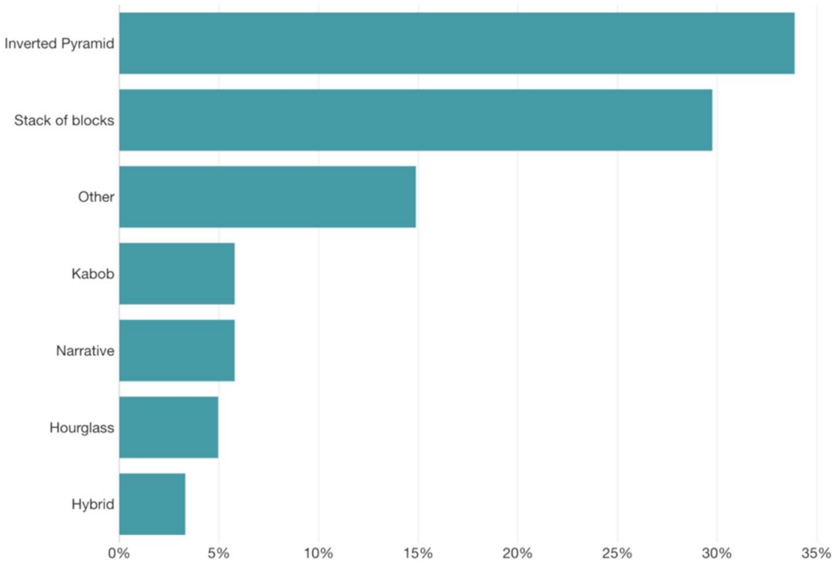


Fig. 10 Storytelling structures used in data stories (N = 118)

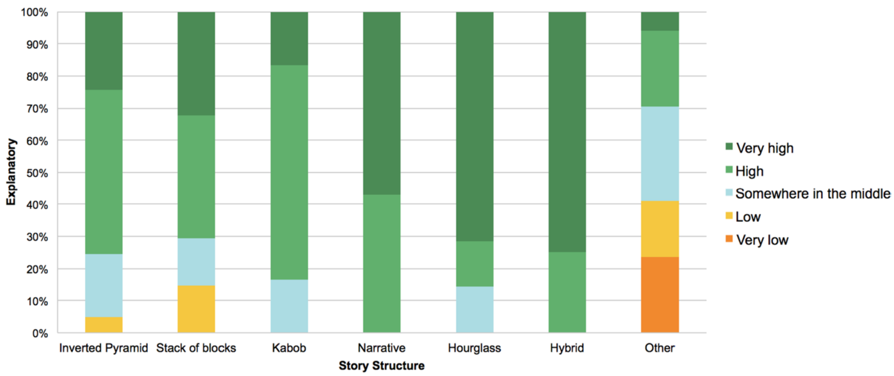


Fig. 11 Story structure and explanatory (H = 29, *p* value = 6.66e−5, Kruskal–Wallis test)

a beginning, middle and end—a dramatic arc. It further supports the claim by Thudt et al., which suggests that "highly explanatory stories are often finite: they have a beginning, an end, and often a linear progression" ([56], p. 68). These are also highly author-driven stories.

On the other end of the spectrum, the Stack of Blocks structure stands out as the structure that presents a notable, yet minimal (only around 10%), portion of stories with low explanatory features.

The stories that did not fall under any of the traditional storytelling structures—tagged as Other—are the only stories that present very low degrees of explanatory characteristics.

Moving onto the examination of exploratory facets, the results indicate a significant difference among various story structures with respect to exploratory characteristics (H = 51, $p < 0.001$, Kruskal–Wallis test). A p value of 3.59e−9 provides additional evidence to support the existence of a significant relationship between story structure and exploratory traits (H = 51, $p < 0.001$, Kruskal–Wallis test). Figure 12 depicts the relationship between story structure and the degrees at which they offer exploration potentials.

Standing out amongst storytelling structures is the Inverted Pyramid structure, where a majority of the stories (nearly 60%) present no exploratory features whatsoever, with another 30% presenting low or very low degrees of exploratory. In other words, 90% of stories in the Inverted Pyramid structure present no or little exploratory features. Similarly, we can also observe that stories with narrative structure offer minimal exploratory features. Yet they offer more exploration compared to the Inverted Pyramid style.

The Stack of Blocks and Hourglass structures are the only two traditional structures that present a notable portion of stories (over 50%) with any degree of explorability at all, i.e. anything other than ‘not at all’ or ‘very low’. Yet, the majority of data stories in the Hourglass structure are ‘somewhere in the middle’ of the exploratory spectrum, with no stories at high or very high degrees of exploratory.

Nearly 20% of data stories in the Stack of Blocks structure exhibit high or very high degrees of exploration, and another 25% of stories in this structure fall within an intermediate range on the exploratory spectrum, indicating moderate degrees of exploratory potential. Putting this next to the explanatory facets of data stories in the Stack of Blocks structure, I would argue that the Stack of Blocks is the only traditional structure that deviates from the high explanatory and low exploratory characteristics of data stories. This could allow for more flexible and versatile data stories.

Only 15% of data stories that did not fall under any of the traditional storytelling structures—tagged as Other—presented high or very high degrees of exploratory features. This may suggest that traditional storytelling structures are to a large extent not versatile enough for data storytelling.

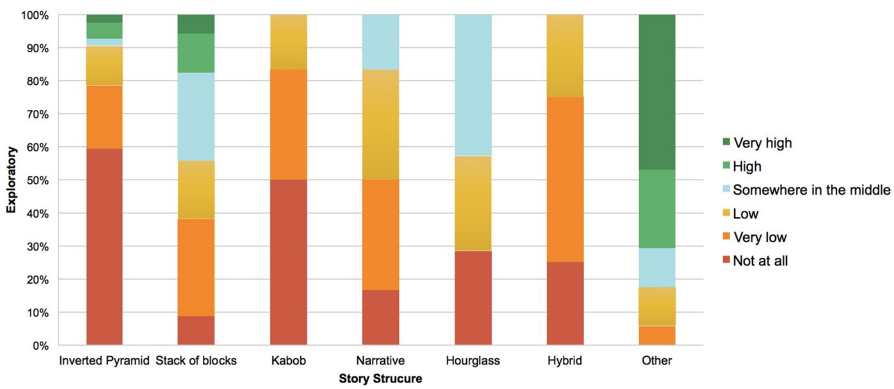


Fig. 12 Story structure and exploratory (H = 51, p value = 3.59e−9, Kruskal–Wallis test)

5 Discussion

Data storytelling, particularly within the realm of data journalism, is a powerful instrument that transforms complex data into stories that engage, inform, and inspire readers [7, 13, 15, 26, 43].

Positioning explanatory and exploratory facets as complementary, rather than opposing characteristics, Thudt et al. [56] argue that balancing exploration and explanation can lead to more engaging experiences and persuasive data stories by providing an author's context and suggested interpretation, alongside the opportunity for the reader to explore a story and pose find the answer to their own questions. Through an analysis of an extensive set of data stories from a diverse range of news outlets, this paper (1) examines the validity of Thudt et al. [56]'s hypothesis regarding the complementary nature of exploratory and explanatory facets, and (2) further investigates the relationship between these two facets and other relevant characteristics of data stories, namely linearity, interactivity, and storytelling structure.

The results indicated that while explanations are predominantly present in data stories, exploratory features also play an increasingly complementary role in enriching data stories in the context of journalism (RQ1). The findings reveal that these two facets do not belong to two ends of a linear spectrum; rather, they can work in tandem, effectively creating a two-dimensional space where both contribute significantly to the story. As such the results support Thudt et al. [56]'s positioning, and by extension, their argument that the interplay between explanatory and exploratory can mutually enhance the storytelling experience. This dynamic interaction allows for richer stories, providing both structured information and opportunities for reader interaction. Such an approach can make data stories more engaging and accessible, catering to diverse audience preferences.

In terms of other characteristics, Linearity in stories was significantly associated with their exploratory and explanatory natures (RQ2). Stories rich in exploratory potentials were found to be more reader-driven, enabling a higher degree of audience control and participatory engagement. On the other hand, stories with a pronounced explanatory presence were typically more author-driven, reflecting a stronger sense of control and direction from the author's perspective.

The study finds a nuanced interplay between the degree of interactivity in data stories and their exploratory and explanatory characteristics (RQ3). The results indicated that data stories with higher exploratory potentials are more likely to be interactive, promoting greater reader engagement. However, it also emerged that interactivity is not exclusive to exploratory stories; even those stories that are more explanatory in nature occasionally incorporate interactive elements.

The analysis of storytelling structures unveiled an intricate relationship with the exploratory and explanatory aspects of data stories. The more traditional storytelling forms, such as the Narrative style and Inverted Pyramid structure, were found to largely exhibit explanatory traits with minimal exploration potential. Specifically, the infrequent application of the Narrative style in data storytelling, notwithstanding its traditional standing, indicates a potential misalignment with the evolving demands of this field.

The Stack of Blocks structure emerged as a more balanced representation of both facets compared to other traditional structures. This suggests it lends itself particularly well to the complementary nature of explanatory and exploratory facets. As one of the most frequently utilised structures in data stories studies, the Stack of Blocks presents a fitting adaptation to the demands of data storytelling amongst traditional storytelling structures, but it also harbours the potential to serve effectively as a baseline for the development of new storytelling structures specific to data stories. One such structure is the Water Tower structure, proposed by Heravt [27, 25]. The Water Tower structure, presented in Fig. 13, integrates the two most commonly used structures in data journalism by starting with an Inverted Pyramid style that presents key information upfront and then transitioning into a Stack of Blocks style for the body, allowing for deeper exploration of various relevant topics. This structure maintains reader interest through the use of 'gold coins' such as data visualisations, providing a balanced and engaging approach to data storytelling.

The introduction of innovative structures, exemplified by the Water Tower model [27, 25], offers promising avenues for the future of data storytelling. This model Such models, which amalgamate elements from traditional narrative forms, could cater more aptly to the dynamic nature of data-driven journalism. At the same time, given the free and innovative nature of many data stories, there will always be stories that would not fit into any box in terms of storytelling structures.

A particularly noteworthy subset of data stories (n = 17), which did not conform to any traditional structure, were also the ones presenting high degrees of exploratory features. Most of these data stories included elements of gamification of the topic. An example is The Pudding's 'The Big Data of Big Hair' story [44], in which

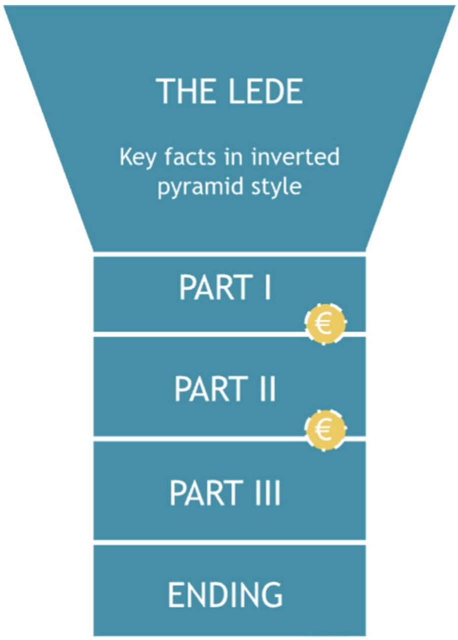


Fig. 13 The Water Tower structure: A new structure for storytelling with data proposed by Heravi [27, 25]

the user can move between dates between 1930 and 2013 and get a set of pictures of women and men depicting their hairstyles (Fig. 14). Another example in this list is the FiveThirtyEight story on ‘Which 2020 Contender Has The Widest Appeal? You Tell Us’ [12]. In this story, the reader guesses the level of support each 2020 US election democratic contender had each of five major constituencies (defined as Party Loyalists, The Left, Millennials and Friends, Black voters and Hispanic and Asian voters). The reader could change each of these characteristics for each contender, and see the appeal ranking change accordingly (Fig. 15).

These stories, which could be seen as innovative, boundary-pushing or even artistic in nature, underscore the need for more flexible storytelling models in data journalism. Some other stories in this group were those that provided primarily search and drill-down functions to the users, such as ProPublica’s Nonprofit Explorer, OjoPublico’s Funes and QuenEsQien.wiki.

This research brings clarity to the intricate dynamics of data storytelling in journalism. It highlights the rising importance and potential for innovation in this field, underscoring the need for flexible, adaptable storytelling structures that can meet the evolving demands of data journalism. The findings from this study shed light on how these features could be used in conjunction with crafting data stories that are both engaging and informative, catering for a wide range of audiences. As such, it

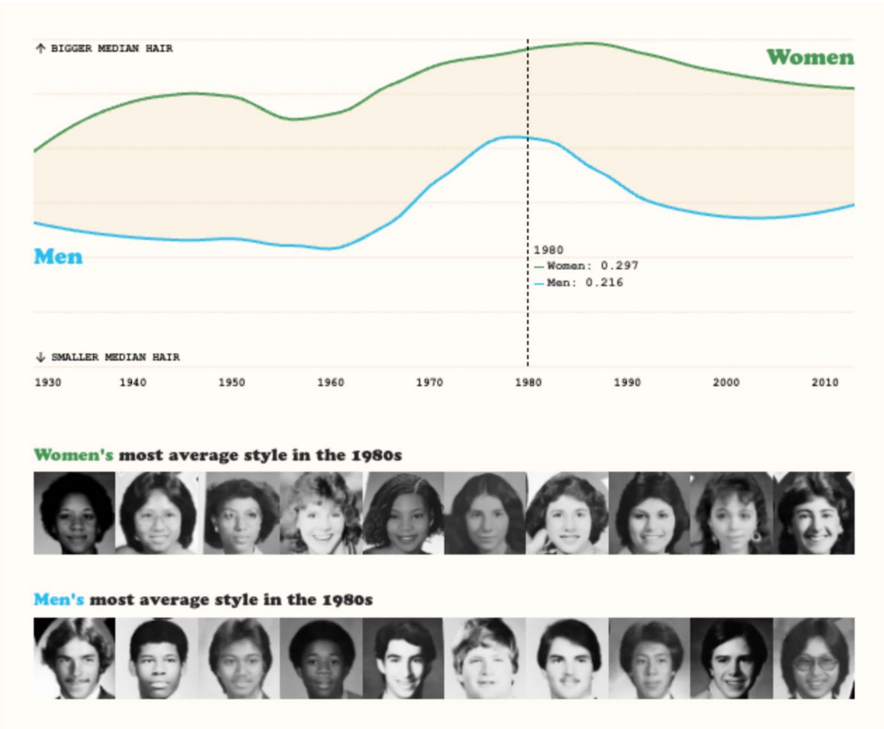


Fig. 14 A screenshot of The Pudding’s The Big Data of Big Hair data story [44]

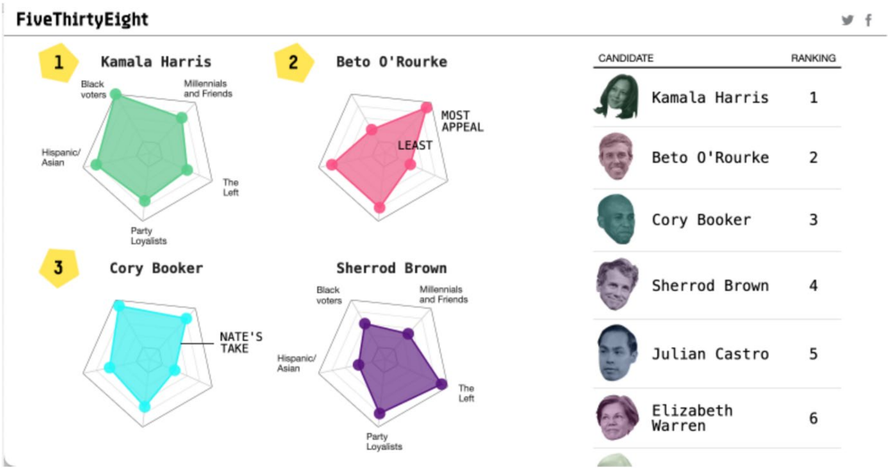


Fig. 15 A screenshot of FiveThirtyEigh 'Which 2020 Contender Has The Widest Appeal? You Tell Us.' data story [12]

not only advances the field of journalism but also sets the stage for broader applications in other disciplines where data storytelling plays a crucial role.

Previous research has examined the relationship between interactivity or storytelling structure with interest and comprehension [2], empathy [21], information processing and memorability [34], engagement and understanding [33] amongst others. According to Thudt et al. [56], we can expect that the integration of exploratory and explanatory features, along with an in-depth understanding of associated characteristics, allows for richer story construction, catering to diverse reader preferences and enhancing overall engagement. Explanatory and exploratory features, and their complementary nature in the context of data storytelling, however, have not been studied in the literature for their impact in this regard. Although such studies remain outside of the scope of this paper, the impact of inclusion or exclusion of each of these characteristics in data stories of these dynamics from industry and audience perspectives, i.e. level of engagement, understanding, empathy, medium, etc. extend beyond the context of this paper, warrant further exploration.

Computational approaches to automated data storytelling are still in their infancy, but they are gaining increasing traction in both research and practice. Studies such as the present one can contribute meaningfully to this emerging area by clarifying the editorial and structural dimensions of storytelling that can be computational models often overlook. While relatively few studies have addressed this space in depth, a small number of works offer initial frameworks. For example, Wang et al. [59] analyse a large corpus of infographics to identify structural and compositional patterns, proposing a system for generating fact sheets via three phases: fact extraction, composition, and presentation. Similarly, Shi et al. [51] introduce a system that organises facts into a logical sequence and presents them as interactive visual stories. These models offer valuable insights into how data stories might be partially automated, however, they do not yet address the structural or editorial aspects of storytelling, particularly at the level of presentation. Other recent research, such as Ye et al. [62], explores the use of Generative AI in the automated creation of data visualisations. However, they primarily studies focus primarily on the technical integration of GenAI into visualisation tools, rather than on the narrative or structural aspects of data storytelling.

The present study complements this body of work by providing detailed empirical grounding for those aspects of storytelling that remain under-specified, particularly in relation to story structure, audience agency, and communicative intent. These findings offer not only a contribution to journalism studies but also valuable input for the development of hybrid, (semi-) automated storytelling systems that better reflect editorial practice and user engagement.

Building on the empirical findings of this study and the wider context of existing literature, the following sections present the core contributions of the paper: first, a set of theoretical advancements that reframe how structure, interactivity, and narrative intent function in data storytelling; and second, a set of practice-oriented recommendations for both journalistic and computational applications.

5.1 Theoretical advancements

This paper advances the theoretical understanding of data storytelling in six key ways:

A reusable Analytical Framework for Data Stories: The study introduces an analytical framework for examining key characteristics of data stories—including explanatory and exploratory dimensions, interactivity, linearity, and storytelling structure. Developed through empirical analysis, the framework offers a structured approach for comparative research and can be adapted for use in related studies across journalism, communication, and computational storytelling.

Re-conceptualising the Exploration and Explanation features: This study supports framing exploration and explanation as independent yet complementary dimensions rather than oppositional. Instead of existing at opposite ends of a linear spectrum, the two can co-occur across a two-dimensional space. This challenges traditional dichotomies and supports a two-dimensional model where data stories can simultaneously exhibit high levels of both explanatory clarity and exploratory depth, offering a more flexible model for understanding how data stories can be constructed to serve varying purposes.

Narrative control and reader agency as a duality: The relationship between explanatory and author-driven characteristics, versus exploratory and reader-driven characteristics, underscores a duality in narrative control. This duality is entangled with the story's structure and interactivity, offering a framework for understanding how data storytellers negotiate control and agency. This suggests that narrative control and reader agency are not binary choices but exist in a dynamic tension shaped by structure, interactivity, and editorial intent.

Integration of storytelling structures: The study provides empirical support for the modularity of story structures, particularly the Stack of Blocks and the

Water Tower structures. These structures facilitate the blending of explanatory and exploratory elements, offering a flexible approach to data storytelling that aligns better with both journalistic and computational models.

Development of specialised structures for Data Storytelling: The findings of this study can shed light on unique needs of data stories, and contribute towards the development of new storytelling structures specifically suited to the affordances and demands of data storytelling and data journalism. One such structure is the Water Tower structure [27, 25], which was conceptually grounded in the empirical work underpinning in this study. These advancements position the Water Tower model not only as a practical tool for storytellers but also as a theoretical framework that encapsulates the dynamic interplay of narrative elements in data-driven contexts.

AI-driven and Computational Storytelling models: The modular nature of storytelling structures such as the Stack of Blocks and the Water Tower could also support (semi-) automated story generation using Generative or rule-based AI systems and computational workflows, or hybrid human–AI systems. For instance, an explanatory lead might be generated using AI-driven summarisation, while subsequent modules could be produced through retrieval-augmented generation based on audience queries, datasets, or structured prompts. This provides a concrete way to integrate editorial values and narrative intent into generative outputs—a key gap in current automated storytelling. In this sense, the structural insights developed in this study bridge empirical journalism research with the computational frameworks proposed in recent work on data storytelling (e.g., [14, 17, 51, 59, 61, 62]), and offer pathways for aligning data storytelling with emerging AI technologies.

5.2 Practical recommendations

Drawing upon the empirical findings, the following academically grounded, practice-oriented recommendations are proposed:

1. **Hybrid formats for balancing Explanation and Exploration:** This study found that stories combining explanatory and exploratory features were not only common but often employed structured formats that enabled both editorial guidance and audience-driven exploration. These formats typically present key messages early, followed by opportunities for modular or interactive engagement. While not all stories require hybrid forms—and many are best served by purely explanatory structures—the repeated use of these patterns across diverse outlets suggests they are an effective strategy for balancing narrative intent with reader agency.
2. **Interactivity as a narrative device, not a design add-on:** Interactivity should not be treated as an add-on or aesthetic feature. Our analysis shows that interactive elements were most effective when aligned with the story's explanatory or exploratory goals—whether guiding the reader or opening space for discovery. This finding supports earlier studies, which argue that interactivity must be anchored in narrative logic to enhance comprehension and engagement. Newsrooms should

therefore embed interactivity within story flows, rather than layering it onto pre-existing structures.

3. **Aligning Story Structures with editorial intent:** Not all stories benefit from the same structure. Our findings show that explanatory stories often follow more linear formats (e.g., Inverted Pyramid), while exploratory stories are better served by modular or interactive designs (e.g., Stack of Blocks or Water Tower). This underscores the value of structure as a narrative affordance. The recommendation is that editors consciously match story goals to structure—treating form not as a template but as a strategic decision.
4. **Supporting modular, cross-functional story development:** While this study did not directly investigate newsroom production processes, the complexity of the stories analysed—particularly those combining explanatory and exploratory elements—suggests the importance of integrated, modular workflows. It observes that many of the most effective stories relied on structural flexibility and interactive elements that would typically require collaboration between editorial, design, and technical teams. Future work may explore how such workflows and tools influence the integration of hybrid storytelling forms in practice.
5. **Data storytelling structures for AI-Assisted storytelling:** Researchers and tool designers working on AI-assisted or automated storytelling should consider adopting modular data storytelling structures—such as the Water Tower— as scaffolding for hybrid systems that combine rule-based logic with generative models. The Water Tower's sequencing, which begins with an explanatory lead followed by distinct modular blocks, lends itself well to computational and AI-assisted decomposition, allowing generative systems to operate within a predefined editorial structure. This hybrid approach retains editorial control while enabling scalable, semi-automated storytelling. The model's modularity and clarity support structured content generation while preserving flexibility in output, making it well suited for workflows that require both narrative coherence and adaptability. In this way, the Water Tower offers a practical bridge between journalistic practice and the design of hybrid storytelling systems. This recommendation also responds to concerns discussed earlier in relation to recent computational models of data storytelling by offering empirically grounded storytelling structures that could inform or be integrated into generative or semi-automated systems.

These recommendations bridge theoretical insights with practical applications, guiding the evolution of data storytelling in journalistic and technical domains.

6 Conclusion

This paper provides a solid foundation and corpus for future research on these topics, as well as for examining how these facets influence audience comprehension and engagement. As the need for effective data stories continues to grow across various areas and mediums, future research should explore these trends and the development of novel and adaptable storytelling structures.

The findings in this paper offer significant insights into the theoretical and practical dimensions of data storytelling. It is important to note here that the relationship analyses in this paper examine the strength of the relationship between the variables, and do not make any causal claims. Additionally, due to the necessity of comprehending the content, the dataset primarily includes stories from English-speaking regions, notably the US and UK. Consequently, the findings of this study may not be universally applicable across different languages and regions, which remains a limitation of this research, and a prompt for future studies to explore how cultural and regional contexts shape data storytelling practices across diverse media environments.

This study opens several other promising directions for future research. One important avenue is examining how different combinations of storytelling structures, explanatory and exploratory elements, and interactivity influence audience engagement, comprehension, and trust. Another is the potential for computational modelling: the structural patterns and editorial dynamics identified here could inform the development of rule-based or hybrid generative systems for semi-automated data storytelling. Relatedly, there is scope for design-oriented work exploring how modular storytelling formats, such as the Water Tower, might be embedded into AI-assisted authoring tools, and how such tools can support editorial control, narrative coherence, and user experience in journalistic workflows.

7 Data availability statement

The datasets curated and tagged for this study will be made available in a public repository of the University of Surrey or OSF, upon acceptance and final publication of the manuscript.

Appendix 1:

Breakdown of publishers included in the dataset.

Publisher	Count	Publisher	Count
Alaska Public Media	1	New Jersey 101.5	1
Athens Live (on Medium)	1	New York TImes	4
BBC	8	OjoPúblico	1
Berkshire Live	1	PODER	1
Bureau Local	5	ProPublica	8
CBC	1	Rappler	1
centralmaine.com	1	Reuters	1
Chicago Tribue	1	Sky News	4
DISCLOSE	1	South China Morning Post	1
El Universal	1	Sunderland Echo	1
Elle Magazine	1	Texas Tribune	1

Publisher	Count	Publisher	Count
Financial Times	4	TEXTY.org.ua	1
FiveThirtyEight	8	The Economist Graphic Detail	6
FT Graphics	17	The Marshall Project	1
Guardian data	8	The Pudding	12
Hereford Times	1	Times Free Press	1
Independent	1	Troika	1
ITV News	1	Upshot	4
Liverpool Echo	1	USA TODAY	1
Naples Daily News	1	Washington Post	3
Grand Total	118		

Declarations

Conflict of interest There is no conflict of interest.

References

1. Al-Rawi, A., Al-Musalli, A., & Fakida, A. (2021). News values on Instagram: A comparative study of international news. *Journalism and Media, 2*(2), 305–320. https://doi.org/10.3390/journalmedia2020018
2. Ameseder, C. A. (2020). Effects of narrative journalism on interest and comprehension: An overview. *Journal of Education and Humanities, 3*(1). https://doi.org/10.14706/JEH2019223
3. Anderson, B., & Borges-Rey, E. (2019). Encoding the UX: User interface as a site of encounter between data journalists and their constructed audiences. *Digital Journalism, 7*(9), 1253–1269. https://doi.org/10.1080/21670811.2019.1607520
4. Appelgren, E. (2018). An illusion of Interactivity: The paternalistic side of data journalism. *Journalism Practice, 12*(3), 308–325. https://doi.org/10.1080/17512786.2017.1299032
5. Bach, B., Stefaner, M., Boy, J., Drucker, S., Bartram, L., Wood, J., Ciuccarelli, P., & Engelhardt, Y. (2018). Narrative design patterns for data-driven storytelling. In N. Henry Riche, C. Hurter, N. Diakopoulos, & S. Carpendale (Eds.), *Data-driven storytelling*. CRC Press.
6. Barlow, M. (2014). *Data visualization: A new language for storytelling*. O'Reilly.
7. Borges-Rey, E., Heravi, B., & Uskali, T. (2018). Periodismo de datos iberoamericano: Desarrollo, contestación y cambio social. Presentación. Revista ICONO14 Revista Científica de Comunicación y Tecnologías Emergentes, 16(2), 1–13. https://doi.org/10.7195/ri14.v16i2.1221
8. Bradshaw, P. (2017). No, I'm not abandoning the term "storytelling", Alberto—Just the opposite (and here's why). Online Journalism Blog. https://onlinejournalismblog.com/2017/09/14/narrative-storytelling-data-journalism-alberto-cairo/

9. Cheshire, J., & Uberti, O. (2014). London: The information capital : 100 maps and graphics that will change how you view the city. Particular Books, an imprint of Penguin Books.
10. Crawford, C. (2003). Assumptions underlying the Erasmatron storytelling system. In P. Sengers & M. Mateas (Eds.), *Narrative intelligence* (pp. 189–197). J. Benjamins Pub.
11. Dick, M. (2014). Interactive infographics and news values. *Digital Journalism, 2*(4), 490–506. https://doi.org/10.1080/21670811.2013.841368
12. Dottle, R. (2019). Which 2020 contender has the widest appeal? You tell us. FiveThirtyEight. https://projects.fivethirtyeight.com/democratic-candidate-appeal/
13. Dykes, B. (2020). *Effective data storytelling: How to drive change with data, narrative and visuals*. Wiley.
14. El Outa, F., Marcel, P., Peralta, V., da Silva, R., Chagnoux, M., & Vassiliadis, P. (2022). Data narrative crafting via a comprehensive and well-founded process. In Chiusano, S., Cerquitelli, T., Wrembel, R. (Eds.), *Advances in databases and information systems. ADBIS 2022. Lecture Notes in Computer Science* (Vol. 13389). Springer, Cham. https://doi.org/10.1007/978-3-031-15740-0_25
15. Emde, K., Klimmt, C., & Schluetz, D. M. (2016). Does storytelling help adolescents to process the news?: A comparison of narrative news and the inverted pyramid. *Journalism Studies, 17*(5), 608–627. https://doi.org/10.1080/1461670X.2015.1006900
16. Fry, D. (2004). Unmuddling Middles. Poynter. https://www.poynter.org/archive/2004/unmuddling-middles/
17. Garretón, M., Moyano, D., Guerra, J., & Parra, D. (2024). Survey of data stories: Guidelines for data story authoring. *Information Visualization, 24*(2), 117–134. https://doi.org/10.1177/14738716241287116
18. Garz, M., & Ots, M. (2025). Media consolidation and news content quality. *Journal of Communication*. https://doi.org/10.1093/joc/jqae053
19. Giles, B. (2004). Thinking about storytelling and narrative journalism. *Nieman Reports, 58*(1), 3.
20. González Veira, X. (2017). I like data-driven storytelling (I like loaded words). Xocas: Data, Visual, Interactive Storytelling. https://medium.com/xocas/i-like-data-driven-storytelling-i-like-loaded-words-d066d83ee909
21. Hand, S., & Varan, D. (2009). Interactive stories and the audience: Why empathy is important. *Computers in Entertainment, 7*(3), 1–14. https://doi.org/10.1145/1594943.1594951
22. Harrower, T. (2013). *Inside reporting: A practical guide to the craft of Journalism* (3rd ed.). McGraw-Hill.
23. Hedstrom, M., & Lee, C. A. (2002). Significant properties of digital objects: Definitions, applications, implications. In *Proceedings of the DLM-forum 2002 access and preservation of electronic information: Best practices and solutions* (pp. 218–227).
24. Heravi, B., Cassidy, K., Davis, E., & Harrower, N. (2021). Preserving data journalism: A systematic literature review. *Journalism Practice*. https://doi.org/10.1080/17512786.2021.1903972
25. Heravi, B. (2024). Storytelling structures in data journalism: Introducing the water tower structure. In S. A. Eldridge & II S. A., Cheruiyot, D., Banjac, S., Swart, J. (Eds.), *Routledge companion to digital journalism studies* (2nd ed., pp. 276–288). Routledge.
26. Heravi, B. (2017). The state of data journalism globally. In *The first European data and computational journalism conference, Dublin, Ireland*. http://researchrepository.ucd.ie/handle/10197/8634
27. Heravi, B. (2022). Storytelling structures in data journalism: Introducing the water tower structure. Computation+Journalism 2022, New York, USA. https://www.academia.edu/84989598/Storytelling_Structures_in_Data_Journalism_Introducing_the_Water_Tower_structure
28. Johnston, J., & Graham, C. (2012). THE NEW, OLD JOURNALISM: Narrative writing in contemporary newspapers. *Journalism Studies, 13*(4), 517–533. https://doi.org/10.1080/1461670X.2011.629803
29. Kapidzic, S., Neuberger, C., Frey, F., Stieglitz, S., & Mirbabaie, M. (2022). How news websites refer to Twitter: A content analysis of Twitter sources in journalism. *Journalism Studies, 23*(10), 1247–1268. https://doi.org/10.1080/1461670X.2022.2078400
30. Knaflic, C. N. (2015). *Storytelling with data: A data visualization guide for business professionals*. Wiley.
31. Knight, M. (2015). Data journalism in the UK: A preliminary analysis of form and content. *Journal of Media Practice, 16*(March), 1–18. https://doi.org/10.1080/14682753.2015.1015801
32. Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). Sage.

33. Kulkarni, S., Thomas, R., Komorowski, M., & Lewis, J. (2023). Innovating online journalism: New ways of storytelling. *Journalism Practice, 17*(9), 1845–1863. https://doi.org/10.1080/17512786.2021.2020675
34. Lang, A. (1989). Effects of chronological presentation of information on processing and memory for broadcast news. *Journal of Broadcasting and Electronic Media, 33*(4), 441–452. https://doi.org/10.1080/08838158909364093
35. Loosen, W., Reimer, J., & De Silva-Schmidt, F. (2017). Data-driven reporting: An on-going (r)evolution? In *An analysis of projects nominated for the Data Journalism Awards 2013–2016*. https://doi.org/10.1177/1464884917735691
36. Machill, M., Köhler, S., & Waldhauser, M. (2007). The use of narrative structures in television news: An experiment in innovative forms of journalistic presentation. *European Journal of Communication, 22*(2), 185–205. https://doi.org/10.1177/0267323107076769
37. McCandless, D. (2009). *Information is beautiful*. Collins.
38. McCandless, D. (2014). *Knowledge is beautiful*. New York: William Collins.
39. Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.
40. Murray, J. H. (2017). *Hamlet on the holodeck: The future of narrative in cyberspace* (Updated). New York: The MIT Press.
41. Van Der Nat, R., Müller, E., & Bakker, P. (2021). Navigating interactive story spaces. The architecture of interactive narratives in online journalism. *Digital Journalism*. https://doi.org/10.1080/21670811.2021.1960178
42. Neveu, E. (2014). Revisiting narrative journalism as one of the futures of journalism. *Journalism Studies, 15*(5), 533–542. https://doi.org/10.1080/1461670X.2014.885683
43. 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
44. O'Brien, E., & Diehm, J. (2019). The big data of big hair. The Pudding. https://pudding.cool/2019/11/big-hair/
45. Pöttker, H. (2003). News and its communicative quality: The inverted pyramid—when and why did it appear? *Journalism Studies, 4*(4), 501–511. https://doi.org/10.1080/1461670032000136596
46. Rogers, S. (2013). *Facts are sacred: The power of data*. Guardian Books.
47. Ryan, M.-L. (2006). *Avatars of story*. University of Minnesota Press.
48. Schröder, J.T., Brück, J. & Guenther, L. (2025). Identifying trust cues: How trust in science is mediated in content about science. *Journal of Science Communication*, 24(1). https://jcom.sissa.it/article/pubid/JCOM_2401_2025_A06/
49. Schulmeister, R. (2003). Taxonomy of multimedia component interactivity. A contribution to the current metadata debate. 61–80.
50. 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
51. Shi, D., Xu, X., Sun, F., Shi, Y., & Cao, N. (2021). Calliope: Automatic visual data story generation from a spreadsheet. *IEEE Transactions on Visualization and Computer Graphics, 27*(2), 453–463. https://doi.org/10.1109/TVCG.2020.3030403
52. Stalph, F. (2017). Classifying data journalism: A content analysis of daily data-driven stories. *Journalism Practice*. https://doi.org/10.1080/17512786.2017.1386583
53. Stalph, F., & Heravi, B. (2021). Exploring data visualisations: An analytical framework based on dimensional components of data artefacts in journalism. *Digital Journalism*. https://doi.org/10.1080/21670811.2021.1957965
54. 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/10.1080/21670811.2015.1042982
55. Tandoc, E. C., Jr., Thomas, R. J., & Bishop, L. (2021). What is (fake) news? Analyzing news values (and more) in fake stories. *Media and Communication, 9*(1), 110–119. https://doi.org/10.17645/mac.v9i1.3331
56. Thudt, A., Walny, J., Gschwandtner, T., Dykes, J., & Stasko, J. (2018). Exploration and explanation in data-driven storytelling. In N. Henry Riche, C. Hurter, N. Diakopoulos, & S. Carpendale (Eds.), *Data-driven storytelling.* CRC Press.
57. Tidwell, J., Brewer, C., & Valencia, A. (2020). *Designing interfaces: Patterns for effective interaction design* (3rd ed.). O'Reilly.
58. Usher, N. (2016). *Interactives journalism: Hackers, data, and code*. University of Illinois Press.

59. Wang, Y., Sun, Z., Zhang, H., Cui, W., Xu, K., Ma, X., & Zhang, D. (2020). DataShot: Automatic generation of fact sheets from tabular data. *IEEE Transactions on Visualization and Computer Graphics, 26*(1), 895–905. https://doi.org/10.1109/TVCG.2019.2934398
60. Weber, W., Engebretsen, M., & Kennedy, H. (2018). Data stories. Rethinking journalistic storytelling in the context of data journalism. *Studies in Communication Sciences*. https://doi.org/10.24434/j.scoms.2018.01.013
61. Yang, Y., Wang, Z., Kim, H., & Lee, B. (2022). A design space for applying the Freytag's pyramid structure to data stories. *IEEE Transactions on Visualization and Computer Graphics, 28*(1), 922–932. https://doi.org/10.1109/TVCG.2021.3114774
62. Ye, Y., Hao, J., Hou, Y., Wang, Z., Xiao, S., Luo, Y., & Zeng, W. (2024). Generative AI for visualization: State of the art and future directions. *Visual Informatics, 8*(2), 43–66. https://doi.org/10.1016/j.visinf.2024.04.003
63. 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/10.1080/17512786.2016.1270171
64. Ytreberg, E. (2001). Moving out of the inverted pyramid: Narratives and descriptions in television news. *Journalism Studies, 2*(3), 357–371. https://doi.org/10.1080/14616700118194