Design Patterns for Data-Driven News Articles

Shan Hao
Shanghai University
Shanghai, China
hshan112592@gmail.com

Zezhong Wang
Simon Fraser University
Vancouver, Canada
wangzezhong2016@gmail.com

Benjamin Bach
University of Edinburgh
Edinburgh, UK
bbach@ed.ac.uk

Larissa Pschetz
University of Edinburgh
Edinburgh, UK
L.Pschetz@ed.ac.uk

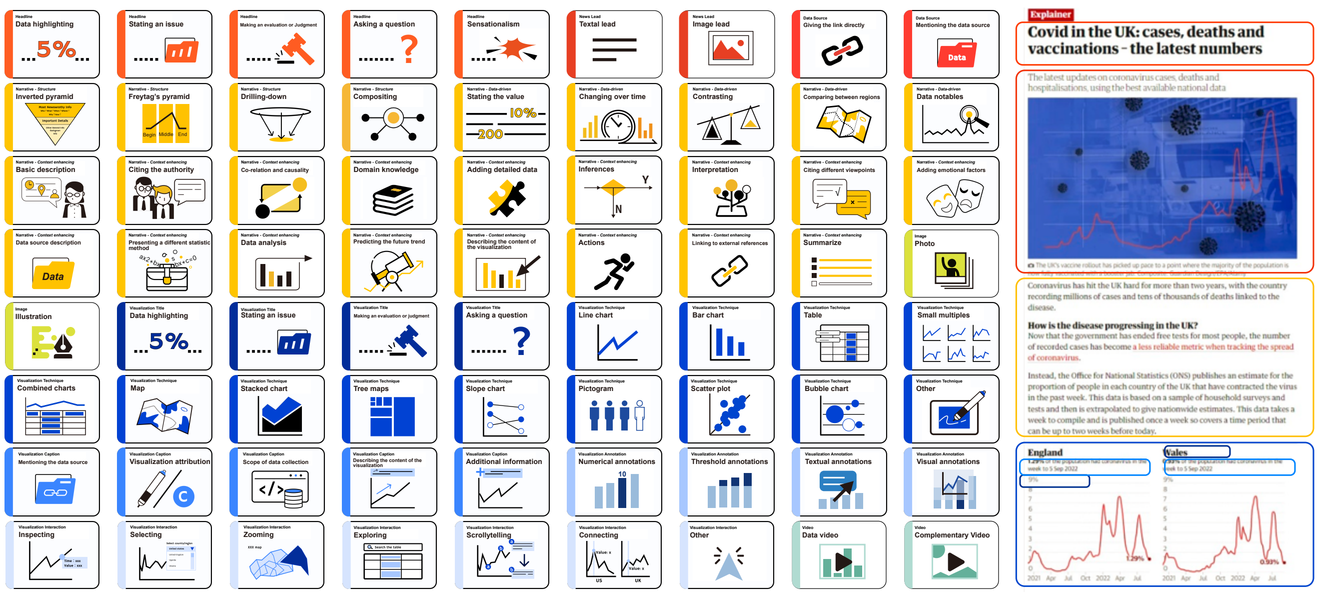

Figure 1: Design patterns (left) and the corresponding use identified from a data-driven news article example (right).

ABSTRACT

Technological advancements have resulted in great shifts in the production and consumption of news articles. This, in turn, lead to the requirement of new educational and practical frameworks. In this paper, we present a classification of data-driven news articles and related design patterns defined to describe their visual and textual components. Through the analysis of 162 data-driven news articles collected from news media, we identified five types of articles based on the level of data involvement and narrative complexity: Quick Update, Briefing, Chart Description, Investigation, and In-depth Investigation. We then identified 72 design patterns to understand and construct data-driven news articles. To evaluate this approach, we conducted workshops with 23 students from journalism, design, and sociology who were newly introduced to the subject. Our findings suggest that our approach can be used as an out-of-box framework for the formulation of plans and consideration of details in the workflow of data-driven news creation.

CCS CONCEPTS

• **Human-centered computing** → **Information visualization**; • **General and reference** → **Design**.

KEYWORDS

Design Patterns, Data Journalism, Data-Driven Storytelling, Education, Classification

ACM Reference Format:
Shan Hao, Zezhong Wang, Benjamin Bach, and Larissa Pschetz. 2024. Design Patterns for Data-Driven News Articles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3613904.3641916

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3641916

1 INTRODUCTION

Data journalism integrates traditional journalism with analytics approaches like data gathering, cleaning, analysis, visualization, and publication [31]. Data-driven news articles are artefacts that blend the textual narrative with data and visualizations and have recently become a pillar in the contemporary news landscape. Currently, the creation of such articles is now shaped by the emergence of digital technologies and new methods of dissemination, accommodating both detailed and selective reading habits, for instance, by emphasizing key excerpts [1]. Furthermore, the landscape of article production and consumption is undergoing profound shifts due to the swift expansion of data resources, a diverse array of data mining technologies, the establishment of communal databases, and worldwide collaborations [26, 49]. Such advancements have given rise to what can be termed as 'big data articles', marking a significant evolution in the field [10, 38]. Eventually, there are many ways to design data-driven articles, e.g., by integrating dashboards [4] and interactive visualizations, leveraging storytelling and other methods from data-driven storytelling [19, 59].

Because of this, designing rich and informative data-driven articles presents a complex challenge; Creators are required to navigate a multitude of design decisions, encompassing genre conventions, data visualization, interactivity, depth of analysis and explanation of visualization and data, including strategies such as highlighting the main messages in the visualizations. While existing research, notably on visual elements and interaction techniques [74], presentation frameworks [43, 53], and key features [69] of data-driven new articles tend to focus on long-form stories or award-winning projects. Shorter articles and the more daily routine practices have not received adequate exploration. Hence, many educators and novices still do grapple with communicative skills for effectively integrating data visualization and news articles. A lack of understanding of these components and their combination makes it hard to make deliberate design decisions.

In this paper, we provide a taxonomy of article types alongside a set of design patterns to describe the richness of data-driven news articles and inform creators' design choices. By analyzing articles from 6 major news media websites, we were able to classify data-driven news articles into five categories: *Quick updates*, *Briefings*, *Chart Descriptions*, *Investigations*, and *In-Depth Investigations*. This framework is intended to facilitate determining the communication goal, the scope of information, and the emphasis when composing an article. Next, we identified 72 design patterns across 11 groups, each pattern describing common components and solutions for data-driven news articles. Besides their analytical value, these design patterns can serve as an educational framework for structuring and presenting information in articles in a modular and scalable design approach.

Eventually, we conducted interviews with 7 experts and educators in the field of data-driven storytelling and journalism, and evaluated the design patterns in workshops with 23 participants to create outlines for data-driven news articles. The outcomes demonstrated that identifying the article types and design patterns can assist participants in understanding the nuance of structure and opens up a broader spectrum of creative opportunities for data-driven news article construction. A detailed description of all our design patterns, and the workshop can be accessed online at: https://datadrivenarticle.github.io/

2 RELATED WORK

A data-driven news article is a data journalistic artefact that presents a *journalistic data story* [69] in a *magazine style* (in the taxonmy by Segel and Heer [59]). In the following, we review literature about classification, creation process, and creation support for data-driven news articles.

2.1 Classification of Data Journalism

With the dawn of the digital age, data journalism has witnessed a transformative journey. It began with the advent of *computer-assisted reporting* in the 1950s [23, 37], gained momentum with the implementation of the Freedom of Information Act and related legislation from 1960s [58], conceptualized in 2008 [27], and has now flourished into a commonplace practice of data journalism [10]. Global news organizations have undertaken extensive data journalism projects with variations in their structure and use of visualization [68]. Faced with this diverse landscape of practices, describing types of data-driven articles can benefit teaching and comprehension of this form of news [57], help meet the diverse reading preferences [1], assist creators in selecting projects aligned with their interests and capabilities [68], and can offer templates and structured references for recurring practical projects [48].

Several taxonomies have been developed to describe the characteristics and production practices in data journalism. For example, Uskali et al. [67] sub-classify data-driven journalism into *investigative data journalism*, *general data journalism*, and *real-time data journalism*, based on criteria such as production time, story scope, technical level, data set, data validation, and analysis method. Similarly, Borges-Rey [9] describes data journalism practice in the UK as either a) daily, with a quick turnaround, generally-visualized, and brief, b) extensive, thoroughly researched and investigative; or c) light, editorialized, entertaining, often-humorous, and gamified. Google News Lab's classification [57] focuses on the relation between story and data and describes three types—*Stories Enriched by Data* in which data is used to verify the report; *Stories Using Data to Investigate* which surfaces stories hidden in the data; and *Stories Explaining Data* that focus on providing meaning behind the data. All these classifications help establish a high-level understanding of data journalism and its production. Yet, they pay limited attention to the *form* of data-driven news articles, their content, their the construction, and the use of data visualization. Our classification of article types and their design patterns fills this gap by providing specific solutions, ideas, and templates.

2.2 Data-Driven Story Creation Process

Several models describe the creation process of data-driven stories. For example, Kosara et al. [39] develop a basic model based on journalistic practices from information collection to presentation. Lee et al. [42] describe three main components of data-driven storytelling process—exploring data, making a story, and telling a story. Building on this work, Chevalier et al. [16] describe artifacts and roles involved in the storytelling process and note that crafting a story and constructing story material require different skill sets.

Similarly, the production process in data journalism is described as *getting data*, *understanding data*, and *delivering data* [10], and by Mirko Lorenz [12] as four stages—*compile* (i.e., start with either a question that needs data, or a dataset that needs questioning), *clean*, *context*, and *combine*. This creation process is rarely always linear and journalists adopt a multi-faceted, iterative, and often cyclical process to news creation [61]. Our work primarily concentrates on the stage of narrative construction, providing materials for teaching, and designing data-driven news articles, rather than offering technical support for specialized tasks within those stages, such as data analysis or creating data visualization.

2.3 Data-Driven News Article Creation Support

Educational materials and creation support for data journalism include handbooks with the introduction of journalistic practice in newsrooms with case studies [10, 27], models of narrative structures of news articles (e.g., [18, 33]), and the key elements in data journalism [18, 69, 74]. Regarding specific approaches for constructing journalistic stories, McKane [46] identifies the *in medias res* approach, which starts with the most dramatic moment and then moves on to the background and less important details (also known as the inverted pyramid structure [50]). Another option is the chronological approach which is rooted in Aristotle's Tragedy Structure: the beginning, middle, and end [22]. This basic structure is further extended by Freytag [24] as *Exposition, Rising action, Climax, Falling action, and Resolution*, has found application in data stories and journalism [20, 71]. Other models provide various strategies for organizing the content in investigative news articles, such as the *Martini Glass* [59], the *Kabob* [32], the *Stack of Blocks* [25], and the *Water Tower* [33].

In addition to structure, design guidance introduced in general data stories and visualizations can serve as valuable inspiration for crafting data-driven news articles. Examples include genres of narrative visualization [47, 59], structures for sequencing [35], the interplay between text and visualizations [41], or story synthesis from visual analytics [15]. Some studies have further supported the education and creation of data storytelling by describing design space and patterns, such as narrative patterns for data-driven storytelling [6], design patterns for data comics [7], narrative semantics of Data Videos [60], dashboard [4], embellishment [14] and composition [36] for data visualizations. The creation supports mentioned above provide invaluable insights, including a high-level workflow, guidance for individual components and specific techniques, and specific tools or automation techniques. However, as these high-level theoretical categories fall short in offering practical guidance, a gap remains for novices in the creation of data-driven news articles. Our approach provides a practical approach, intended to assist creators in informing the structure of data-driven news articles and to bridge the gap between theoretical knowledge and practical use, particularly for novices to data-driven articles.

3 METHODOLOGY

Our research methods included 1) article sampling, iterative coding of 2) article types and 3) design patterns, as well as user evaluation through 4) one-on-one sessions and 5) workshops (Fig. 2).

3.1 Article Sampling and Article Coding

Our initial sample of articles focused on the global outbreak of COVID-19 (Stage 1, Fig. 2). As an international emergency, COVID-19-related news reports provided a confined topic, lots of data reporting, and allowed comparison among those articles. We conducted our sampling from 6 prominent English-language news media, chosen for their high website traffic, dedicated data columns, diverse editorial styles, and global reputation for representative journalism. All of these media outlets cater to a wide range of reader demographics [63] and political inclinations [30], providing a rich and varied sample of data-driven news articles. The data-driven news articles were sampled from The Guardian (n = 26), The Economist (n = 20), The Times (n = 17), Financial Times (n = 18), New York Times (n = 12), and BBC News (n = 9), and were randomly sampled from the respective COVID-19 related columns. In the process of sampling, we ensured that the collected data-driven news articles were presented in magazine style [59] , and contained at least one data visualization each. All 102 samples were published between March 2020 and April 2022.

Through coding these articles into different types, we aimed to understand how the evolving production and consumption landscape influences the presentation of these articles. This aids in comprehending their structural layout and positioning within the broader landscape. Our approach was inspired by the work of Uskali [67] and Rogers [57], who categorized various data parameters including data sources, quantity, journalistic verification, production cycle, and usage. However, our focus was primarily on the types of articles rather than the process of their creation. As such, we adapted their classification method by omitting elements like production cycle and technical level.

We classified articles by variation of five criteria (detailed in Section 4) and employed both inductive and open coding methods (Stage 2, Fig. 2). Two coders coded the same 15 articles independently and improved the codes based on three rounds of discussions until an agreement was reached. Then, one of the authors completed the coding of all 102 articles. The coding resulted in labels describing the level of five criteria by describing how each criterion is presented in an article, which was then used to classify data-driven news articles into 5 types (Section 4).

3.2 Coding for Design Patterns

In this second coding stage (Stage 2) we coded articles according to common design pattern for an article's syntactic components and their visualizations: headline, data sources, narratives, visualization titles, visualization techniques, visualization annotations, visualization captions and interactions. To prevent uniformity, we extended the set of articles for this coding stage to topics other than COVID-19. First, we randomly chose 4 articles about COVID-19 articles from each article type from our initial set of 102 articles. Second, we went back to the 6 media outlets and sampled new data-driven news articles from the respective columns on economy, society, culture, politics, health, and others (Fig. 3). However, through random sampling, we found that many articles belong to our type *investigations* and hence do not sufficiently cover our other four article types. In the hope of increasing article type coverage, we searched for other articles by individual authors named on these randomly

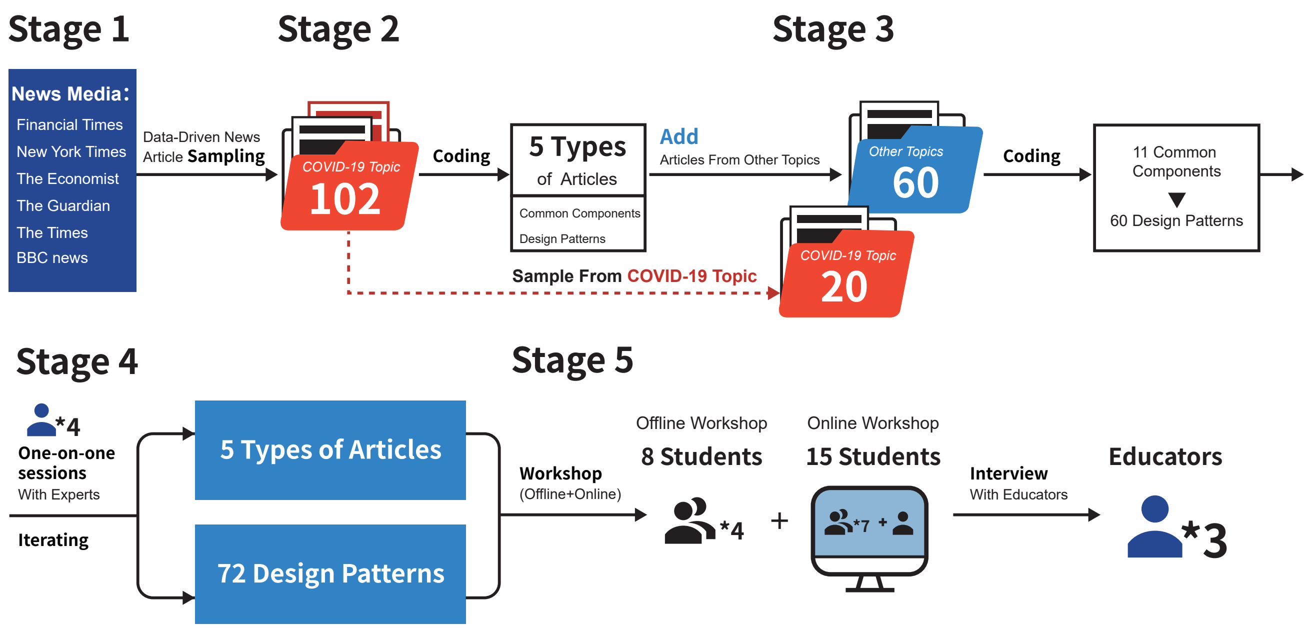

Figure 2: Overview of study methods. Data-driven news articles were sampled and coded twice, once for article types (Stage 2) and again for design patterns (Stage 3). Stage 4 involves soliciting feedback for improvements from experts. In Stage 5, we conducted workshops and interviewed educators to evaluate the use of our design patterns.

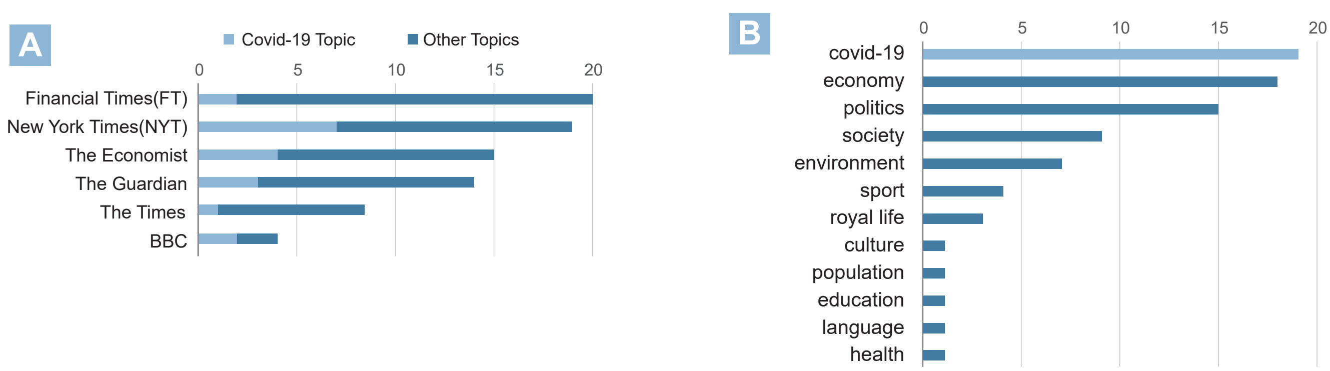

Figure 3: Frequency of (A) sources and (B) topics in the 80 data-driven articles sampled in Stage 3 (Fig.2)

sampled articles. Our assumption was that these individual articles might be shorter while still published by the same media outlet. These individual authors identified their roles as 'visual project editor,' 'data project, visual storytelling reporter,' and 'information designer.' In total, this gave us 35 articles. Together with 25 of the randomly sampled team-authored articles, and the 20 articles from the COVID-19 topic, our sample set for Stage 3 coding was 80. Last, we coded all of those new 60 articles according to our article types (Stage 3) to verify we did not encounter any new article types. All those articles were published between 2019 to 2022 and followed the same selection criteria as our initial set of 102 articles. Across all 80 articles, our samples contained 1,224 paragraphs and 313 visualizations, corresponding to 15.3 paragraphs and 3.9 visualizations per article.

The same person who coded all 102 COVID-19 corpus coded design patterns for each of these components and engaged in discussions with the other co-authors to refine the codes. Once saturation in refinement was reached, the same person coded all 80 articles. In total, we identified 60 design patterns (Figure 1, details in Section 5).

3.3 Iteration and Evaluation

To obtain initial feedback on the classification of the article types and design patterns (Stage 4, Fig. 2), we conducted one-on-one sessions with one data journalist (14 years experience as a professional journalist) and three data-driven storytelling and visualization researchers (2 PhD students and 1 post-doctoral researcher). In the one-on-one sessions, we described the type of article to the expert and showed relevant article cases, as well as our design pattern cards, then invited the expert to use the cards to construct a simple article framework, which helped the expert discover problems and provide feedback. Afterward, we asked the expert for their experience in design articles, how our approach could be used in their workflow, and potential improvements. According to the feedback, we modified some names and descriptions of our article types and patterns to enhance clarity. For instance, the article type that provides summary information was renamed from *Daily News* to *Briefing* to avoid giving audiences the impression that it was solely based on information from the same day. We added narrative structures from existing literature [18, 33] to serve as inspiration for narrative construction. We have also added a few design patterns based on their suggestions. For example, *Inference* was split from the original coding as part of the *interpretation.* After refining the codes we recorded all articles again. Lastly, we evaluated the design patterns in multiple workshops with students from journalism, and design backgrounds and interviewed their lectures about potential future usage of involving our design patterns in their classrooms (Stage 5, Fig. 2).

4 FIVE TYPES OF DATA-DRIVEN NEWS ARTICLES

We identified five types of data-driven news articles in Stage 2 of our process. We categorized articles based on five key criteria:

- **C1: Data**—The **range of data sources** from single to multiple, collection techniques from public and open access data sets or unofficial information sources and self-collected data, and analytical methods.
- **C2: Analysis**—The **depth of data description**, ranging from overview summaries to detailed explanations, and whether visualizations are employed to support the data.
- **C3: Context**—The degree of **contextual detail** provided, from minimal to extensive.
- **C4: Investigation**—The **breadth and depth of the investigation**, topic, and factors associated with the event, from narrow to wide-ranging.
- **C5: Discussion**—The **depth of discussion**, from superficial to comprehensive.

We specifically focus on content and structure, as these components indicate the depth of data analysis, the logic used, and the intended objectives of data-driven news articles. Our focus does not extend to the authoring process in terms of the collaboration model among journalists, the number of collaborators, technical proficiency, or production time. We also abstain from categorizing articles based on text length or the number of visualizations. For example, *Magazine Dashboards* [4], could be longer if they display multi-faceted information with many visualizations or shorter if they focus on a single aspect; however, only the length and numbers of data visualizations do not necessarily dictate a change in the article type. Also, the five types are not intended to have rigid boundaries; rather, they serve to describe five typical representations along a spectrum.

4.1 Article Type: Quick Updates

Quick Updates generally offer a snapshot of current data, featuring visualizations as the key elements and limiting textual descriptions for quick comprehension. Similar to Uskali's *Real-Time News* [67], data presented in the articles are automatically updated, and *Magazine Dashboards* [4] (Fig. 4A, [N1]), which provides at-a-glance insight [73], *Quick Updates* is commonly used in journalism for presenting data related to public health or political elections, e.g., data update during the COVID-19 pandemic period [75]. According to the examples we coded, *Quick Updates* mainly use open public data (C1), and cover multiple facets of an event with limited text descriptions (C2), the main component of this article type is visualizations.

4.2 Article Type: Briefings

Briefings provides concise summaries and overviews of events. It typically includes a summarized overview of an event and data over a certain period, offering audiences a glimpse of trends, or the latest developments. For example, the article from BBC News titled *Coronavirus: What is the R number and how is it calculated?* (Fig. 4B, [B4]) presents a brief explanation of the R value, its variations across different regions in the UK, and three key infectious disease metrics. The article provides a high-level overview of the subject without delving into detailed interpretations. Similarly, the World This Week column in The Economist, featuring concise articles like *Business* [E2], offers varied economic summaries without delving deeply into analysis.

Briefings typically source their data from open data or data from other published articles (C1). They provide event summaries and highlight significant data changes but do not delve deep into the reasons behind the data (C2). More than just presenting minimum descriptions as in *Quick Updates*, the textual descriptions in *Briefings* are more informative, and visualizations are used as supplements rather than being main components. In *Briefings*, the depth of data analysis remains shallow, primarily offering a collection of summary-level insights (C3).

4.3 Article Type: Chart Descriptions

Chart Descriptions indicates news articles with an explicit narrative structure, depicts the development of events with analysis of more than one aspect. Regarding data sources, *Chart Descriptions* typically uses publicly available data to focus on various facets of events (C1). While they offer observations, they provide only limited insights (C2). There is ample relevance between the context and the description of data, the text supplements the causes and consequences of the development of the event (C3).

This type is akin to what Rogers referred to as *stories that are enriched by data* [57], leaning more towards traditional reporting, where data is used to validate foundational reporting. The example from The New York Times titled *What the BA.5 Subvariant Could Mean for the United States* (Fig. 4C, [NYT7]) explores the changes brought about by the BA.5 subvariant of COVID-19 in the United States, offering a multifaceted explanation of the challenges associated with this variant. The article then goes on to predict the future trajectory of cases in the United States by drawing comparisons with the situations in other countries.

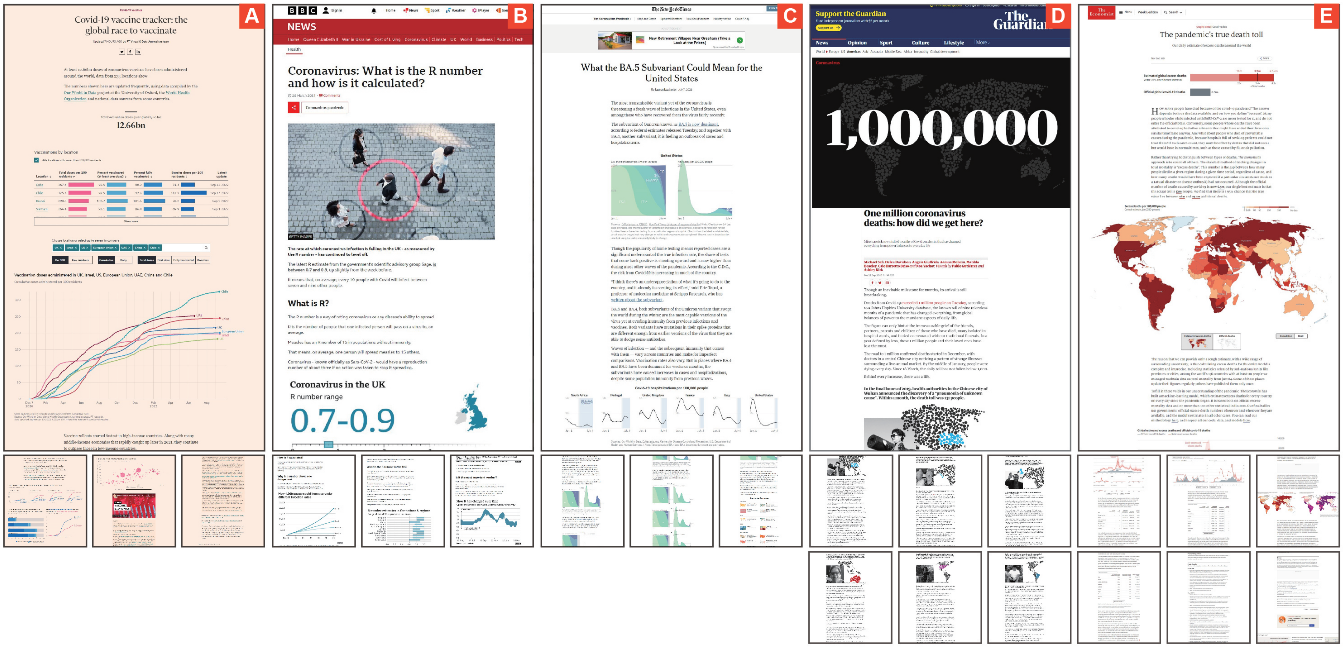

Figure 4: Examples of the five types (Section 4) of data-driven news articles, with thumbnails showing the continuation of the articles above. (A) Quick Update example: COVID-19 vaccine tracker: the global race to vaccinate. (by the Financial Times, [FT2]); (B) Briefing example: Coronavirus: What is the R number and how is it calculated? (by BBC News, [B4]); (C) Chart Description example: What the BA.5 Subvariant Could Mean for the United States. (by the New York Times, [NYT7]); (D) Investigation example: One million coronavirus deaths: how did we get here? (by the Guardian, [28]); (E) In-Depth Investigation example: The pandemic's true death toll (by The Economist [21])

4.4 Article Type: Investigations

Investigations aims to uncover critical issues or provide a deeper understanding of complex phenomena. This type of article draws on a range of data sources, including government open datasets and academic reports, and may also include data unearthed by journalists, to provide a well-rounded view of the issue (C1). It often involves storytelling through data-driven narratives using visualizations with interactions to engage and inform audiences (C2). These articles exhibit close contextual connections, and each section builds upon the previous one in a progressive manner (C3). Journalists compile and analyze fragments of information from diverse sources to reveal hidden patterns, trends, and insights accompanied by diverse types of data visualization techniques (C4), to guide readers to reflect on the event (C5). The Guardian's article titled *One million coronavirus deaths: how did we get here?* (Fig. 4D, [28]) chronicles the progression of COVID-19, leading to one million deaths in a span of nine months. Through the creation of character sketches of individuals who succumbed to the virus in various countries and interviews with their family members, along with an exploration of the overall development of the disease, the story sheds light on the resultant policy and lifestyle changes, and the grief experienced by the loved ones of the deceased. By investigating the personal stories behind the death toll, the report encourages readers to contemplate more deeply the real-life impact of the relevant data.

4.5 Article Type: In-Depth Investigations

The *In-Depth Investigations* features articles that delve much deeper than surface-level reporting, utilizing comprehensive data analyses to uncover complex trends, relationships, or anomalies. These pieces often involve extensive research, rich data sources from multiple channels (C1), and sophisticated statistical and data visualization techniques (C2), to bring new insights through lengthy narratives (C3). The aim is to provide a thorough understanding of the issue at hand (C4), often challenging traditional wisdom or revealing previously hidden aspects of a subject, thereby fostering critical discussion and thought (C5).

An example is The Economist's article titled *The pandemic's true death toll* (Fig. 4E [21]). The authors expressed skepticism regarding official reports of death tolls, prompting them to collect data from 121 indicators from over 200 countries and regions and create a machine-learning model. According to the estimation, the authors provide a central figure of 10.2 million deaths, which was twice the official estimate and even as low as one-quarter in some cases. Using new data sources and estimation models, the article offers a different perspective with novel insights.

5 DESIGN PATTERNS FOR DATA-DRIVEN NEWS ARTICLES

5.1 Design Pattern Descriptions

This section describes 72 design patterns for data-driven news articles (Figure 5), identified by coding the 80 articles in Stage 3, (Figure 2). Patterns are categorized according to 11 article components: *Headline, News Lead, Data Source, Narrative, Image, Visualization Title, Visualization Technique, Visualization Annotation, Visualization Caption, Visualization Interaction,* and *Video*. Percentages indicate the frequency of articles applying a given pattern. Percentages for each component do not sum up to 100% because multiple patterns can be used in combination. Percentages associated with design patterns for visualization titles, visualization techniques, visualization annotations, visualization captions, and visualization interactions indicate pattern occurrence *across all articles, per article*, not pattern frequency per visualization. In other words, a 10% value means 10% of our 80 articles feature at least one visualization with the respective pattern.

Headline Patterns—Prior studies have explored the pivotal role of headlines in news articles, along with the sentiments they convey [3, 17]. However, research targeting headlines in data-driven news articles remains notably limited. According to our coding, patterns for headline ranges from directly referring data to different forms of summary. ***Data highlighting*** (3.75%) directly specifies numerical values in the headline. ***Stating an issue*** (40%) not directly displaying data but using dataset or event names, as in G4 'Covid in the UK: Cases, deaths…'. ***Making an evaluation or judgment*** (19.75%) states a finding from the data or an identified issue. ***Asking a question*** (36.25%) formulates questions as headlines, similar to *an open framework* [8]. ***Sensationalism*** (1.25%) employs exaggerated and dramatic language to capture attention.

News Lead Patterns—The news lead, also referred to as the 'news lede', is situated below the headline and concisely introduces audiences to the main points and essential information of the article [18]. ***Textual lead*** (71.25%) introduces the article's content through text, and ***Image lead*** (63.75%) uses images, photographs, or a combination of visuals and data visualizations (56.86%) to represent the article's content.

Data Source Patterns—The data source is a critical factor that influences the credibility and transparency of a news article. Data sources were specified by ***Giving the link directly*** (73.75%) (e.g., URL) to the database, or ***Mentioning the data source*** (25%) in the text. We found that only one article in our samples does not specify the data source.

Narrative Patterns—Regarding narrative framework, we drew inspiration from previous research, including the analysis of data tasks [2], the recognition of semantic content within articles as explored [45], and the fine-grained analysis of text narrative presented by Latif et al. [41], who analyzed 22 data-driven geographic news stories, distinguishing between data-driven and embedding text narratives. Our encoding builds upon their work, presenting an extended version and diversity introduced by various article and visualization types. We categorize text narratives into three distinct types: structure, data-driven text, and context-enhancing text, we have compiled 26 relevant design patterns for these narrative types.

Previous research has explored narrative structures in news articles, although not exclusively those that are data-driven (e.g., [18, 33]). Four patterns are commonly used in our samples. ***Inverted pyramid*** structural pattern [50], commonly employed in traditional journalism and still prevalent in data journalism. This pattern starts with the most important information and then provides background information and conclusions. ***Freytag's pyramid*** pattern [24] consists of five phases in a narrative: Exposition, Rising action, Climax, Falling action, and Resolution. The ***Drilling-down*** approach transitions from a broad overview to a detailed, micro-level analysis, offering an in-depth exploration of particular facets within a larger subject. The ***Compositing*** structure centers on a unifying theme, presenting various pieces of information in either a complementary or parallel fashion to create a comprehensive view. Each section independently explains an aspect of the topic. It may also take the form of a question-and-answer format, where questions are used as titles for subsections.

There are five types of data-driven text patterns: ***Stating the value*** (10%) simply presents the values or the change of the value. ***Change over time*** (63.75%) describes the trend and comparison of values at different times. ***Contrasting*** (75%) is to compare the data, such as contrasting a whole with its parts or present rankings. The ***Comparing between regions*** approach (30%) compares data values across different geographical areas. ***Data notables*** (15%) focuses on specific values such as outliers and extrema.

Context-enhancing text patterns include: ***Basic description*** (62.5%) presents a summary or an overview. ***Citing the authority*** (71.25%) involves referencing government or academic reports, as well as quoting experts to enhance the article's credibility. ***Co-relation and causality*** (68.75%) puts data into context, reveals the co-relation of events, and explains reasons. As in the G12 '…these indicators have important implications for the level of inequality: taxation, social spending in sectors such as health, welfare and education, and labor rights'. To elaborate on co-relation or causality, creators may extend to explaining concepts in specific ***domain knowledge*** (30%). For example, G14 introduces the problem of drinking water quality by specifying the pollutants and their effects, 'EPA stipulates that the nitrate content per liter of water is 10 mg, but it is often exceeded. This standard is designed to prevent the fetus from getting enough oxygen.' ***Adding detailed data*** (72.5%) is to provide specific examples or more in-depth data. ***Inferences*** (60%) is to provide appropriate inferences according to the data. ***Interpretation*** (50%) illustrates links in the development of things, e.g., 'Despite the additional disruption caused by Brexit, developed economies around the world are dealing with similar problems' (G11). ***Citing different viewpoints*** (18.75%) to present different opinions by citing perspectives from different entities. The ***Adding Emotional Factors*** category (constituting 11.25% of articles surveyed) refers to the intentional inclusion of emotionally resonant elements in the article to elicit empathy from readers. For instance, in coverage of the COVID-19 pandemic, an

Figure 5: The full list of our design patterns for data-driven news articles with examples. The patterns are identified by coding 80 articles in Stage 3 (Fig. 2), also informed by literature on narrative structures in journalism, and design patterns such as description of visualization techniques [66] and visualization interaction [11, 65, 72].

article might not only present the statistical increase in cases but also provide real-life examples of individuals who succumbed to the virus. Family interviews and testimonials can serve as emotional touchpoints, adding a human element to the raw data, thereby making the information more relatable and impactful. ***Data source description*** (38.75%) is to provide information about the issuing organization, sample method, and whether the data is preprocessed or filtered for certain reasons. ***Presenting a different statistic method*** (16.25%) involves the use of novel statistical approaches to reveal insights that may not be apparent using conventional methods of data analysis. This category aims to provide fresh perspectives on existing data by employing less commonly used statistical techniques, thereby shedding new light on a topic or issue. ***Data analysis*** (13.75%) is to present the process of analyzing the data, interpreting the statistics, and drawing conclusions. ***Predicting the future trend*** (26.25%) is to predict the future trend of an event based on the current data. ***Describing the content of the visualization*** (25%) involves providing an interpretation or summary of what the data visualization is presented. ***Actions*** (20%) focuses on detailing the actions that have been taken or are recommended to be taken in light of the data-driven findings. **Linking to external references** (85%) is to include links to other articles in the text as references. ***Summarize*** (90%) means to conclude from the pre-mentioned data and arguments as a take-home message of a section or the entire article.

Image Patterns—In our analysis, we identified two image design patterns in the articles: ***Photo***, used in 17.5% of articles (14 out of 80), and ***Illustration***, found in two instances, e.g., illustrations are used as separators between chapters [Ft14].

Visualization Title Patterns have been found to use the same patterns as for article headlines.

Visualization Technique Patterns—Through statistical analysis of 313 visualizations in the 80 articles, we found that ***Line charts*** were the most frequently employed type (26.2%), followed by ***Bar charts*** (13.1%), ***Tables*** (10.22%), ***Small multiples*** (10.22%), ***Combined charts*** (9.27%), ***Maps*** (6.07%), ***Tree maps*** (4.79%), ***Stacked charts*** (3.51%), ***Slope graphs*** (2.24%) and ***Pictogram*** (2.24%). A smaller proportion of visualization techniques are also found such as ***Scatter plots*** (1.6%), ***Bubble charts*** (1.6%) and ***Other*** chart types (8.95%), such as grid charts, heat maps, violin plots, and innovative illustrations.

Visualization Annotation Patterns—There are many types of annotations by their forms. Ren et al. [56] conducted an analysis of variations in line charts, bar charts, and scatter plots in data journalism visualizations, proposing visual annotation methods such as text and shapes. Building upon this, and influenced by the presence of visualizations like maps and information charts in our corpus, our analysis delved further. We identified additional annotation types beyond numerical annotations, leading us to recode and categorize our findings into four types: ***Numerical annotations*** (76.92%) represent the most prevalent category, encompassing numerical value near its visual representation. ***Threshold annotations*** (15.2%) delineate intervals or values with specific significance. ***Textual annotations*** (9.47%) describe data information or provide additional summary-like information. ***Visual annotations*** (14.79%) refer to the supplementary visualization annotated in a visualization.

Visualization Caption Patterns—Visualization caption refers to the textual descriptions that typically appear as footnotes beneath the visualization or occasionally below the title (N=2). Out of 246 charts, we found patterns including ***Mentioning the data source*** (84.5%), ***Visualization attribution*** (60.16%) that states the creator or source of the visualization, ***Scope of data collection*** (48.78%) that refers to the contextual parameters within which data is collected, often specifying aspects like the time range, geographical locations, or any other conditions that define the boundaries for data acquisition. ***Describing the content of the visualization*** (15.85%) is the same as the pattern in the narrative. Lastly, ***Additional information*** (4.47%) adds extra details, such as influencing factors to a trend or contextual information related to the visualization.

Visualization Interaction Patterns—Yi et al. [72] explored interactions for analysis, and subsequently, Boy et al. [11] supplemented Yi's work, including semantic operations such as inspection, which have been adopted in data journalism [74]. In our sample articles, 124 charts (40.13%) employed interactions, revealing eight interaction intents: ***Inspecting*** (70.16%) reveals detailed data on hover. ***Selecting*** (39.51%) enables audiences to opt for specific items to view, often via drop-down menus or by highlighting particular elements. ***Zooming*** (12.1%) enable audiences to display more or less detail. ***Exploring*** (9.68%) allows audiences to search for information according to their preferences. ***Scrollytelling*** (4.48%) navigates audiences to the next section through scrolling. ***Connecting*** (3.23%) links all related elements of a certain element for comparison. And a portion of ***Other*** (6.45%) forms of interaction, such as games or links to anchors of contents within the article.

Video Pattern—Within our corpus, six articles (7.5%) incorporated videos, all of which originated from the Financial Times. We classified videos into ***Data video*** (40%), wherein information is conveyed as animated visual narratives with data visualizations [60]. ***Complementary video*** (60%), providing additional information to enhance the topic beyond the data, such as background information, interviews, and expert opinions.

5.2 Association Between Article Types and Design Pattern Usage

While coding patterns, we sometimes observed trends between article types and design patterns summarized in Figure 6. For instance, *Quick Updates* often use data names as headlines patterns, while *Investigations* and *In-Depth Investigations* used more subjective or abstract headlines (Fig. 6 A). Narrative patterns, we found, vary with the depth of data discussion (C5)(Fig. 6 B),e.g., *Quick Updates*

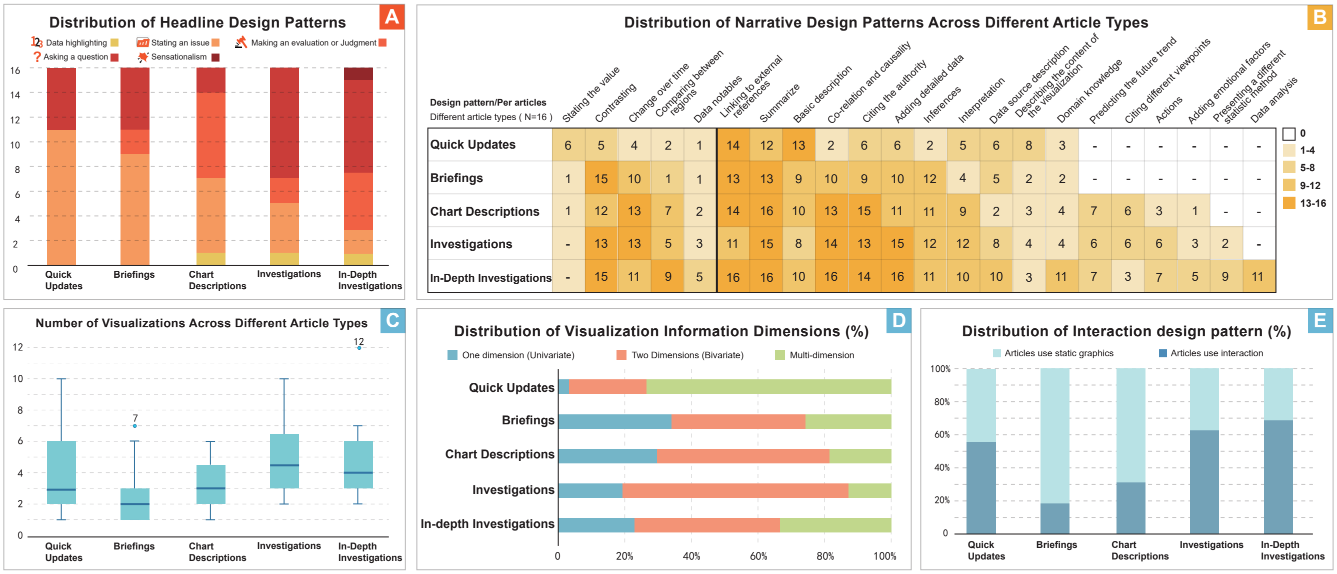

Figure 6: (A) Distribution of Headline design patterns; (B) Distribution of Narrative design patterns; (C) Number of Visualizations; (D) Distribution of Visualization Information Dimensions; (E) Distribution of Interaction design patterns across different article types.

commonly use the Stating-the-value pattern, while *Investigation* articles often employ the Comparing-between-regions and the Data-notables pattern. Context-enhancing patterns like Linking-to-external-references are widespread, enhancing dissemination and visibility. *Chart Descriptions* and *Investigations* frequently employ Citing-the-authority, Adding-detailed-data, and Co-relation-and-causality, while Data-analyis and Presenting-a-different-statistic-method are exclusive to *In-Depth investigations*.

Regarding Visualization patterns, although there is no notable variance in the visualization titles, the quantity and complexity of visualizations vary. *Quick Updates* and *Investigations* include more visualizations, with *Quick Updates* showing higher data density, often revealed through hovering or other interactive means. Visualization Interaction patterns increase from *Briefings* to *In-Depth Investigations* (Fig. 6 C, D, E).

6 WORKSHOP

We conducted a workshop (Stage 5, Fig. 2) to evaluate to if article types and design patterns can help novice journalists understand and create data-driven news articles. The workshop also sought to identify the challenges creators faced throughout the creation process.

6.1 Participants

The workshops were open to students from different backgrounds who were interested in crafting data-driven news articles. It did not require any specific skills. We recruited 23 participants (16 females) aged between 19 and 35 by advertising on online social media platforms and directly contacting course instructors teaching design, digital media, and journalism from three universities in China and the UK. 13 students majored in journalism and communication (1 Ph.D. student, 12 undergraduate students); 8 in visual, interaction design (3 Ph.D. students, 3 master students, 2 undergraduate students); 2 graduate students in law and ecological economics respectively. Seven participants (30%) had previously learned data visualization in a university course and had experience in creating between 3 and 5 data-driven articles. The remaining participants had read articles, but had no practical experience creating them. We conducted one offline workshop (8 people) and 8 online workshops (15 people), and the participants were divided into 13 groups (Stage 5, Fig. 2).

6.2 Data and Materials

We prepared workshop materials around the topic of cancer, which allowed for diverse possible perspectives for varied story types. We provided participants with a *Briefing* article on the topic of 'The 5 types of cancer with the most deaths domestically'. Alongside, we provided data sets on cancer as well as 4-5 data visualizations on the proportion of new cancer cases across countries and national cancer incidences. Participants were free to seek additional data online.

6.3 Procedure

All workshops were run by one of the authors on Miro (https://miro.com) (Stage 5, Fig. 2), accompanied by a different university course educator each. The online workshops were with students from the journalism, law and ecological economics disciplines. The offline workshops were with students from design discipline. After the workshops, we interviewed three of the educators, two in graphic design, one in journalism. Each educator had 3-4 years of university teaching experience. Both online and offline workshops had a duration of 2.5 hours each. We began with a 30-minute introduction to data-driven journalism, data stories, and data visualization, then introduced the five article types and design patterns with examples.

After the introduction, participants were divided into groups of two on a voluntary basis. In the online workshops, participants joined in pairs, with one session featuring one participant. We provided the *Briefing* article and the other materials. Each group was asked to create one type of article chosen from *Chart Descriptions* to *In-Depth Investigations* by using inspiration from the design patterns. For offline workshops, we offered physical copies of the design patterns and worksheets as cards. Each card contains the title, icon, brief description and case of the design pattern. Online workshops featured digital copies of those cards on a shared Miro board. This creation phase took about 60-70 minutes. Participants were told to think aloud and be free to discuss with their team members.

Finally, we asked each group to present their article outline and design rationale. The article outline mimics the structure of a data-driven news article, with design pattern cards being selected and dragged into the corresponding parts of the article. For example, selecting a headline card and placing it at the beginning, accompanied by a brief textual annotation explaining how the title was created using the card prompt. Narrative cards are similarly selected and placed in order, with a short textual description explaining the content. To reflect on their design process and associated challenges, we conducted semi-structured interviews with online workshop participants and questionnaires with offline participants. Questions centered around *Whether our article types and design patterns were helpful for creation, and if so, in what ways? What were the challenges you encountered in your creation process? How can our approach be improved? Did you employ other digital assistants, such as ChatGPT, during the creation process? If so, how did you utilize them?* All sessions of the online workshops were video and audio recorded. Following the workshop, we conducted interviews with three educators. We presented five types of data-driven news articles and related design patterns and engaged in semi-structured interviews that started with two questions: *Whether our approach could assist learning and teaching, and if they would like to use these materials for teaching, how they would implement them?*

6.4 Results

The outcome of all workshops included 13 data-driven article outlines presented by assembled design pattern cards, brief textual descriptions of their design rationales on worksheets, as well as feedback gathered from interviews and questionnaires.

One group chose *In-Depth Investigations* on the topic of medical fees and insurance; they used ChatGPT for data queries. Eight groups created *Investigations*, with the topic falling into two categories: regional comparisons and detailed analyses of particular types of cancer, such as breast and lung cancer. Participants utilized diverse sources for data collection, including datasets we supplied, and enriched their research by querying additional data and materials through search engines, data from academic publications, and ChatGPT. Four groups of students created *Chart Descriptions*, with three groups highlighting the differences in cancer prevalence across regions and among different populations, and one group focusing on the relationship between food temperature and the risk of cancer.

Regarding the article construction process, we noted a variety of approaches to structuring the articles. The design patterns were used in a flexible way rather than following the same workflow. Five groups focused on the data and visualization techniques first, then moved to the narrative section, while the other eight groups discussed the narrative aspect first, followed by the selection of the visualization.

Classification of article types——Based on our observations and feedback, we found that ***identifying the article type initially helps to set expectations for the article's content and guides the scope of initial data searches***. In our workshops, identifying and determining the article type was the first step in the construction process, contributing to the overall planning of the article, and forming an initial selection of the design pattern cards for the framework. Group 8 participant 1 (G8P1): *"The questions of what type you plan to write made it easier for me to make choices on the narrative"*; Educator 2 (Edu2): *"Identifying the article type is an important initial step in article construction, providing students with a directional framework for their work."*).

Design pattern usability——Our design patterns effectively ***enable beginners to rapidly engage in hands-on practice, facilitating the creation of data-driven news articles***. The type of articles and design pattern cards gave participants a rich pool of inspirational options for their writing. Even participants new to data journalism and data visualization can discuss and use the design pattern cards to frame the articles. Among the seven groups who constructed *Investigations*, each group employed 9 to 12 narrative patterns. The most frequently used data-driven text patterns were *Change over time* (by 7 groups) and *Contrasting* pattern (by 6 groups). Additionally, participants made extensive use of context-enhancing patterns, including *Citing the authority* and *Summarize* (by all 7 groups), as well as *Co-relation and causality* and *Domain knowledge* (by 6 groups). Participants' pattern usage closely mirrored the patterns we coded in the *Investigations*, suggesting that participants' planning of the depth and direction of their investigative endeavors aligns with the established conventions of this article type. G7P1: *"The patterns are very comprehensive and work as small molecules. Beginners like me were able to understand and use them to build an article framework"*; Edu1: *"[design patterns] contribute to the development of students' abilities to organize material effectively, which in turn encompasses both the richness of the content and the clarity of its presentation"*.

Due to the variation of article types constructed by participants, it is challenging to identify clear trends in the use of different types and design patterns. However, we still observed some design patterns that aligned with participant preferences. For instance, the use of headline patterns, *Ask a Question* (5 groups) and *Sensationalism* (4 groups) are more frequently used, indicating that participants aimed to create titles that effectively capture attention. Narrative design patterns showed *Contrast* being frequently used (13 groups) in data-driven narratives, followed by *Change over time* (7 groups). In context-enhancing aspects, patterns like *Summarize*, *Citing the authority*, and *Correlation and causality* were frequent (11 groups each). Interestingly, *Action* was also widely used (9 groups).

We also observed that 6 groups used ChatGPT during the creative process. They explored various directions for their inquiries, including data sources (e.g., G1: 'Could you suggest websites that offer digital health reports containing relevant data on cancer?'), understanding a topic (e.g., G2: 'What are the factors that cause cancer?' 'What foods are likely to cause cancer?'), participants used design patterns to frame their questions enquiring ideas for specific components (e.g., G1: 'Provide a headline idea with a *sensational* tone.' and G2: 'What are *co-relation or causality* between lifestyle and cancer?').

Providing diverse methods for presenting information broadens the range of creative possibilities in writing. Integrating data into articles and visualizations was a recurring challenge and focal point in group discussions among the participants. They found our design patterns useful in inspiring them to address this challenge. For example, G4P2: *After learning so many ways of interaction, I have a clearer idea about how to deal with a large amount of data.* The discussion in groups was mainly about making decisions among these options such as visualization techniques and proper interactions to be applied.

Increasing awareness of data transparency. Indicating the sources of data is a demonstration of journalistic transparency [64]. The data source design patterns served as a reminder for participants to emphasize data authenticity and credibility, facilitating their inclusion of these elements in their articles. All of the participants used this pattern in their outlines.

Educators have expressed appreciation for our article types and design patterns, believing that these materials can aid students in their reporting-writing process. Regarding the usability of our materials in teaching, Edu1, the lecturer in journalism, suggests that our materials can be integrated into courses in news reporting and interviewing in journalism. Journalism writing demands both efficiency and objective accuracy, which extends beyond sourcing data or obtaining foundational materials. It also requires efficient textual narrative structures, an area where our materials can provide valuable guidance. For teaching interviews in journalism, *"[narrative pattern cards] offer a variety of options, which can help students organize the logical flow of their writing, thereby avoiding major oversights or logical inconsistencies during the writing or interview process".* Design instructor Edu2 expressed that our materials can assist students in developing effective thinking patterns and practical approaches. In the visualization techniques section, commonly used data visualization techniques are introduced, providing students with valuable guidance for creating visualizations. He recommends, '*expanding the introduction of data visualization design tools to provide further guidance for professionals specializing in data visualization design.*'

7 DISCUSSION

7.1 Creating Data-Driven News Articles with Design Patterns

By analyzing articles from data columns and the personal websites of data journalists, we identify and describe a spectrum of five types of data-driven news articles, ranging from *Quick Updates* to *In-depth Investigations.* Built on the existing classification of data journalism by Google News Lab [57] and Nardelli [55], our article types and design patterns systematically describe the corpus of article structure, and components. This ready-to-use resource is invaluable for teaching, designing, and engaging in discourse on data-driven news articles.

Our design patterns offer several opportunities in the multifaceted process of creating data-driven news articles, which spans data collection, analysis, visualization, and article composition. Firstly, they act as cognitive aids for conceptualizing and structuring data-driven articles, thus enhancing both writing and investigative skills. These patterns offer a rich information pool that helps kickstart creative thinking and foster a logically structured flow. Inspired by the successful use of cards for design thinking [7, 29, 34], we create the multi-card format of design patterns to promote alternative methods of presentation and organization with digital and physical activities, and stimulate fruitful discussions.

Additionally, design patterns can enhance the investigative quality of articles as they guide students in formulating questions and in critically interpreting data (e.g., by *Citing different viewpoints* or surface the process of *Inference*), as well as reduce the likelihood of omitting crucial details (e.g., by *Mentioning the data source*). Essentially, article type and design patterns offer a ready-to-use toolbox for aspiring authors, benefiting both educational and communicative fields.

However, it is important to note that these patterns alone do not guarantee best practices. They are meant to facilitate the article-writing process by offering structured support. However, the eventual quality of the final article depends on numerous factors, such as writing style, data complexity, quality and suitability of visualizations, on the authoring side, as well as logical reasoning, data literacy, and relevance, on the audience side. Future research could focus on exploring how to involve this approach in current journalism education.

7.2 Efficiency Gains and Formulaicity Risks with Design Patterns for Data-Driven News Article Creation

Although Section 5.2 presents some trends of patterns with specific article types, it is crucial to remain cautious of the formalization risks associated with rigidly applying these patterns. Even though we counted, we opted not to disclose the frequency of pattern usage by article type to avoid creating potential biases for participants. Presenting such statistics from the sample could inadvertently establish a perceived 'standard' for how certain patterns should be used for specific article types. This could lead to a 'one-size-fits-all' approach, stifling individualized or innovative design decisions tailored to unique cases and communication goals. In the workshop, we did not observe the fixed modes of combination patterns. Participants presented various creative approaches. Our design patterns served as a toolbox rather than a standard structure. Participants actively discussed their choices of patterns with their collaborators in the creation process.

Balancing creative freedom with structured guidance is, however, a trade-off in support tools for article creation[52]. Future research could delve into the relationships between article types, design patterns, and communication objectives with a larger sample, potentially informing the development of design tools that offer more nuanced guidance.

7.3 Adaptions and Extensions

Our design patterns for data-driven news articles can extend beyond traditional formats to inform the design of other storytelling genres, such as scrollytelling or slideshow articles that also incorporate titles, textual descriptions, visualizations, and references to data sources. Additionally, the design patterns could be adapted to various social media platforms where long-form investigative pieces are distilled into concise teasers to quickly inform and attract specific audiences, including through visually engaging formats such as GIFs [62]. Thus, our design patterns could provide the basis for the development of a flexible framework for content adaptation across platforms that could inform the development of targeted, platform-specific information algorithms.

Furthermore, our design patterns can be combined by using other low-level narrative devices that serve specific intents, such as the 6 detailed time-oriented narrative sequences (i.e., Chronology, Trace-back, Trailer, Recurrence, Halfway-back, and Anchor) [40], or the 18 narrative design patterns for data-driven storytelling [6] such as Convention breaking and Concretize to enrich the narration in news reporting. By adapting to other layouts, our design patterns can be used to create data-driven news articles in other forms of presentation such as posters, data comics, or data videos.

Finally, our design patterns could help human-machine collaborations. We noted that workshop participants utilized various forms of assistance during the article creation process, including using ChatGPT to seek inspiration for diverse facets of the topic and to locate data resources. Within the realm of journalism, technologies can be seen as collaborative human partners, or 'co-creators of journalism.' [51]. However, current automated technologies are limited in interpreting cultural sensitivity or human emotions which are essential factors for investigative types of journalistic reporting [70]. There is indeed a gap between authoring tools, AI-supported tools [44], and AI-generator tools [54]. Recommended approaches for the latter are typically solely based on data features or design guidelines ignoring creators' intent [13]. Hence, significant untapped potential exists for tools that bridge technologies with creative practice. Our design patterns could serve as a modular framework to facilitate more effective interactions between creators and AI-based generation tools.

7.4 Limitations of the Work

We acknowledge that our article samples are primarily from Western English-speaking media. Expanding the dataset to articles in different languages and from different cultures could provide valuable cultural and complementary perspectives. While our sampled corpus showcases a degree of diversity, limitations in sample size and representativeness remain; e.g., our samples represent what Sparks [63] describes as the 'serious press' and 'semi-serious press', characterized by a blend of traditional news and soft news features. However, our study did not include much of the 'serious-popular press' and 'news stand tabloid press', known for their visual emphasis and focus on scandal, sports, and entertainment, and the ones which leans heavily towards scandal and entertainment while retaining some serious news elements. One could think that our sampling initiated with articles on COVID-19 could be a further limitation. However, when we compared design patterns across articles on COVID-19 with those on other topics, we found no significant differences in their frequency and application.

Eventually, while our paper focuses on the structure and components of data-driven news articles, other elements like layout, font, color, and presentation devices warrant further investigation. While article types and design patterns serve as useful aids for content creation, effective teaching and learning in visualization [5] and data-driven news should also incorporate other critical competencies, such as data and visualization literacy as well as ethical approaches to data and journalistic practice. We focus on novice creators planning to create data-driven news articles. Future research could engage with professionals to expand or refine the patterns and usage.

8 CONCLUSION

In this paper, we describe design patterns for data-driven news articles to support literacy and proficiency in writing data-driven news articles. Building on existing classifications in data journalism and analyzing contemporary examples, we also identified five types of data-driven news articles tailored for various scenarios, aiding creators in setting expectations and swiftly constructing a framework for their articles. Subsequently, we approach the article from a holistic perspective, identifying design patterns under eleven common components of data-driven news articles: headline, news lead, data source, narrative, image, visualization title, visualization techniques, visualization annotation, caption, interaction, and video. These design patterns were refined through expert interviews and workshops. Results suggest that our design patterns could guide beginners in quickly getting started and improving their efficiency in constructing data-driven news articles.

REFERENCES

[1] 2023. Reuters Institute and the University of Oxford Issues Report Entitled 'Reuters Institute Digital News Report 2022'.
[2] R. Amar and J. Stasko. 2004. BEST PAPER: A Knowledge Task-Based Framework for Design and Evaluation of Information Visualizations. In *IEEE Symposium on Information Visualization*. 143–150. https://doi.org/10.1109/INFVIS.2004.10
[3] Faheem Aslam, Tahir Mumtaz Awan, Jabir Hussain Syed, Aisha Kashif, and Mahwish Parveen. 2020. Sentiments and emotions evoked by news headlines of coronavirus disease (COVID-19) outbreak. *Humanities & social sciences communications* 7, 1 (2020), 1–9.
[4] Benjamin Bach, Euan Freeman, Alfie Abdul-Rahman, Cagatay Turkay, Saiful Khan, Yulei Fan, and Min Chen. 2023. Dashboard Design Patterns. *IEEE Transactions on Visualization and Computer Graphics* 29, 1 (2023), 342–352. https://doi.org/10.1109/TVCG.2022.3209448
[5] Benjamin Bach, Mandy Keck, Fateme Rajabiyazdi, Tatiana Losev, Isabel Meirelles, Jason Dykes, Robert S Laramee, Mashael AlKadi, Christina Stoiber, Samuel Huron, et al. 2023. Challenges and opportunities in data visualization education: A call to action. *IEEE Transactions on visualization and computer graphics* (2023).
[6] Benjamin Bach, Moritz Stefaner, Jeremy Boy, Steven Drucker, Lyn Bartram, Jo Wood, Paolo Ciuccarelli, Yuri Engelhardt, Ulrike Koeppen, and Barbara Tversky. 2018. Narrative design patterns for data-driven storytelling. In *Data-driven storytelling*. AK Peters/CRC Press, 107–133.
[7] Benjamin Bach, Zezhong Wang, Matteo Farinella, Dave Murray-Rust, and Nathalie Henry Riche. 2018. Design Patterns for Data Comics. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173612
[8] Stefano Bocconi, Frank Nack, and Lynda Hardman. 2005. Supporting the Generation of Argument Structure within Video Sequences. In *Proceedings of the*

Sixteenth ACM Conference on Hypertext and Hypermedia (Salzburg, Austria) *(HYPERTEXT '05)*. Association for Computing Machinery, New York, NY, USA, 75–84. https://doi.org/10.1145/1083356.1083371
[9] Eddy Borges-Rey. 2016. Unravelling Data Journalism: A study of data journalism practice in British newsrooms. *Journalism practice* 10, 7 (2016), 833–843.
[10] Liliana Bounegru, Liliana Bounegru, and Jonathan Grey. 2021. *The data journalism handbook : towards a critical data practice / edited by Liliana Bounegru and Jonathan Gray.* Amsterdam University Press, Amsterdam.
[11] Jeremy Boy, Francoise Detienne, and Jean-Daniel Fekete. 2015. Storytelling in Information Visualizations: Does It Engage Users to Explore Data?. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 1449–1458. https://doi.org/10.1145/2702123.2702452
[12] Paul Bradshaw. 2011. The inverted pyramid of data journalism. *Online Journalism Blog* 7 (2011).
[13] Qing Chen, Shixiong Cao, Jiazhe Wang, and Nan Cao. 2023. How Does Automation Shape the Process of Narrative Visualization: A Survey of Tools. *IEEE Transactions on Visualization and Computer Graphics* (2023), 1–20. https://doi.org/10.1109/TVCG.2023.3261320
[14] Qing Chen, Ziyan Liu, Chengwei Wang, Xingyu Lan, Ying Chen, Siming Chen, and Nan Cao. 2022. VizBelle: A Design Space of Embellishments for Data Visualization. *arXiv.org* (2022).
[15] Siming Chen, Jie Li, Gennady Andrienko, Natalia Andrienko, Yun Wang, Phong H. Nguyen, and Cagatay Turkay. 2020. Supporting Story Synthesis: Bridging the Gap between Visual Analytics and Storytelling. *IEEE Transactions on Visualization and Computer Graphics* 26, 7 (2020), 2499–2516. https://doi.org/10.1109/TVCG.2018.2889054
[16] Fanny Chevalier, Melanie Tory, Bongshin Lee, Jarke van Wijk, Giuseppe Santucci, Marian Dörk, and Jessica Hullman. 2018. From analysis to communication: Supporting the lifecycle of a story. In *Data-Driven Storytelling*. AK Peters/CRC Press, 151–183.
[17] Celeste M Condit, Alex Ferguson, Rachel Kassel, Chitra Thadhani, Holly Catherine Gooding, and Roxanne Parrott. 2001. An exploratory study of the impact of news headlines on genetic determinism. *Science Communication* 22, 4 (2001), 379–395.
[18] Zeyu Dai, Himanshu Taneja, and Ruihong Huang. 2018. Fine-grained structure-based news genre categorization. In *Proceedings of the Workshop Events and Stories in the News 2018*. 61–67.
[19] Ensheng Dong, Hongru Du, and Lauren Gardner. 2020. An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases* 20, 5 (2020), 533–534.
[20] Brent Dykes. 2019. *Effective Data Storytelling - How to Drive Change with Data, Narrative, and Visuals.* John Wiley & Sons, Newark.
[21] The Economist. 2021. The pandemic's true death toll. https://www.economist.com/graphic-detail/coronavirus-excess-deaths-estimates. [last accessed: Mar 2, 2023].
[22] Gerald Frank Else. 1938. Aristotle on the Beauty of Tragedy. *Harvard studies in classical philology* 49 (1938), 179–204.
[23] James S Ettema and Theodore L Glasser. 1998. *Custodians of conscience: Investigative journalism and public virtue.* Columbia University Press.
[24] Gustav Freytag and Elias J. MacEwan. 1895. *Freytag's Technique of the drama : an exposition of dramatic composition and art.* Benjamin Bloom, New York.
[25] D. Fry. 2009. Stack of Blocks. https://donfry.wordpress.com/2009/06/14/stack-of-blocks/
[26] Lucas Galan, Jordan Osserman, Tim Parker, and Matthew Taylor. 2019. How young people consume news and the implications for mainstream media. (2019).
[27] Jonathan Gray, Lucy Chambers, and Liliana Bounegru. 2012. *The Data Journalism Handbook.* O'Reilly Media, Incorporated, Piraí.
[28] The Guardian. [n. d.]. One million coronavirus deaths: how did we get here? https://www.theguardian.com/world/ng-interactive/2020/sep/29/one-million-coronavirus-deaths-how-did-we-get-here-covid. [last accessed: Nov 26, 2023].
[29] Kim Halskov and Peter Dalsgård. 2006. Inspiration card workshops. In *Proceedings of the 6th conference on Designing Interactive systems.* 2–11.
[30] Michael Hameleers. 2019. Partisan Media, Polarized Audiences? A Qualitative Analysis of Online Political News and Responses in the United States, U.K., and The Netherlands. *International journal of public opinion research* 31, 3 (2019), 485–505.
[31] James T Hamilton. 2016. *Democracy's detectives: The economics of investigative journalism.* Harvard University Press.
[32] Timothy Harrower. 2006. *Inside Reporting: A Practical Guide to the Craft of Journalism.* https://api.semanticscholar.org/CorpusID:111049843
[33] Bahareh Heravi. 2022. Storytelling Structures in Data Journalism: Introducing the Water Tower structure. In *Computation+ Journalism 2022*.
[34] Eva Hornecker. 2010. Creative Idea Exploration within the Structure of a Guiding Framework: The Card Brainstorming Game. In *Proceedings of the Fourth International Conference on Tangible, Embedded, and Embodied Interaction* (Cambridge, Massachusetts, USA) *(TEI '10)*. Association for Computing Machinery, New York, NY, USA, 101–108. https://doi.org/10.1145/1709886.1709905
[35] Jessica Hullman, Steven Drucker, Nathalie Henry Riche, Bongshin Lee, Danyel Fisher, and Eytan Adar. 2013. A Deeper Understanding of Sequence in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2406–2415. https://doi.org/10.1109/TVCG.2013.119
[36] Waqas Javed and Niklas Elmqvist. 2012. Exploring the design space of composite visualization. In *2012 IEEE Pacific Visualization Symposium.* 1–8. https://doi.org/10.1109/PacificVis.2012.6183556
[37] Precision Journalism. 1973. A Reporter's Introduction to Social Science Methods.
[38] Nicolas Kayser-Bril, Anastasia Valeeva, and Irina Radchenko. 2016. Transformation of communication processes: Data journalism. *arXiv preprint arXiv:1605.01956* (2016).
[39] Robert Kosara and Jock Mackinlay. 2013. Storytelling: The Next Step for Visualization. *Computer* 46, 5 (2013), 44–50. https://doi.org/10.1109/MC.2013.36
[40] Xingyu Lan, Xinyue Xu, and Nan Cao. 2021. Understanding Narrative Linearity for Telling Expressive Time-Oriented Stories. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 604, 13 pages. https://doi.org/10.1145/3411764.3445344
[41] Shahid Latif, Siming Chen, and Fabian Beck. 2021. A Deeper Understanding of Visualization-Text Interplay in Geographic Data-driven Stories. In *Computer Graphics Forum*, Vol. 40. Wiley Online Library, 311–322.
[42] Bongshin Lee, Nathalie Henry Riche, Petra Isenberg, and Sheelagh Carpendale. 2015. More Than Telling a Story: Transforming Data into Visually Shared Stories. *IEEE Computer Graphics and Applications* 35, 5 (2015), 84–90. https://doi.org/10.1109/MCG.2015.99
[43] Wiebke Loosen, Julius Reimer, and Fenja De Silva-Schmidt. 2020. Data-driven reporting: An on-going (r) evolution? An analysis of projects nominated for the Data Journalism Awards 2013–2016. *Journalism* 21, 9 (2020), 1246–1263.
[44] Min Lu, Chufeng Wang, Joel Lanir, Nanxuan Zhao, Hanspeter Pfister, Daniel Cohen-Or, and Hui Huang. 2020. Exploring Visual Information Flows in Infographics. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3313831.3376263
[45] Alan Lundgard and Arvind Satyanarayan. 2022. Accessible Visualization via Natural Language Descriptions: A Four-Level Model of Semantic Content. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 1073–1083. https://doi.org/10.1109/TVCG.2021.3114770
[46] Anna McKane. 2013. News writing. *News Writing* (2013), 1–192.
[47] S. McKenna, N. Henry Riche, B. Lee, J. Boy, and M. Meyer. 2017. Visual Narrative Flow: Exploring Factors Shaping Data Visualization Story Reading Experiences. *Computer graphics forum* 36, 3, 377–387.
[48] Faten El Outa, Patrick Marcel, Veronika Peralta, and Panos Vassiliadis. 2023. Highlighting the Importance of Intentional Aspects in Data Narrative Crafting Processes. *Information Systems Frontiers* (2023), 1–17.
[49] Anna Podara, Maria Matsiola, Theodora A Maniou, and George Kalliris. 2019. News usage patterns of young adults in the era of Interactive Journalism. *Strategy & Development Review* 9 (2019).
[50] Horst Pöttker. 2003. News and its communicative quality. The «inverted pyramid»—Origin and establishment of a professional standard in journalism: Die «Umgekehrte Pyramide»—Ursprung und Durchsetzung eines journalistischen Standards. *Publizistik* 48 (2003), 414–426.
[51] Alex Primo and Gabriela Zago. 2015. Who And What Do Journalism?: An actor-network perspective. *Digital journalism* 3, 1 (2015), 38–52.
[52] Larissa Pschetz, Koji Yatani, and Darren Edge. 2014. TurningPoint: narrative-driven presentation planning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* 1591–1594.
[53] Lea Püchel and Christian-Mathias Wellbrock. 2021. Examining the digital renewal of news communication: A categorization of presentation modes in digital journalism. *New Media & Society* (2021), 14614448211059488.
[54] Chunyao Qian, Shizhao Sun, Weiwei Cui, Jian-Guang Lou, Haidong Zhang, and Dongmei Zhang. 2021. Retrieve-Then-Adapt: Example-based Automatic Generation for Proportion-related Infographics. *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2021), 443–452. https://doi.org/10.1109/TVCG.2020.3030448
[55] Alastair Reid. 2015. 'Without humanity, data alone is meaningless': Data journalism insights from the Guardian. https://www.journalism.co.uk/news/-without-humanity-data-alone-is-meaningless-data-journalism-insights-from-the-guardian/s2/a564652/. [last accessed: Mar 2, 2023].
[56] Donghao Ren, Matthew Brehmer, Bongshin Lee, Tobias Höllerer, and Eun Kyoung Choe. 2017. ChartAccent: Annotation for data-driven storytelling. In *2017 IEEE Pacific Visualization Symposium (PacificVis)*. 230–239. https://doi.org/10.1109/PACIFICVIS.2017.8031599
[57] Simon Rogers, Jonathan Schwabish, and Danielle Bowers. 2017. Data journalism in 2017: The current state and challenges facing the field today. *Google News Lab* (2017).

[58] Michael Schudson. 2014. Origins of the freedom of information act in the United States. *Transparency in Politics and the Media. Accountability and Open Government. Reuters Institute for the Study of Journalism & IB Tauris, Oxford* (2014), 1–18.

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

[60] Yang Shi, Xingyu Lan, Jingwen Li, Zhaorui Li, and Nan Cao. 2021. Communicating with Motion: A Design Space for Animated Visual Narratives in Data Videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 605, 13 pages. https://doi.org/10.1145/3411764.3445337

[61] Dilruba Showkat and Eric P. S. Baumer. 2021. Where Do Stories Come From? Examining the Exploration Process in Investigative Data Journalism. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2, Article 390 (oct 2021), 31 pages. https://doi.org/10.1145/3479534

[62] Xinhuan Shu, Aoyu Wu, Junxiu Tang, Benjamin Bach, Yingcai Wu, and Huamin Qu. 2020. What makes a data-gif understandable? *IEEE Transactions on Visualization and Computer Graphics* 27, 2 (2020), 1492–1502.

[63] Colin Sparks and John Tulloch. 2000. *Tabloid tales: Global debates over media standards.* Rowman & Littlefield.

[64] Florian Stalph. 2018. Classifying Data Journalism: A content analysis of daily data-driven stories. *Journalism practice* 12, 10 (2018), 1332–1350.

[65] Nicole Sultanum, Fanny Chevalier, Zoya Bylinskii, and Zhicheng Liu. 2021. Leveraging Text-Chart Links to Support Authoring of Data-Driven Articles with VizFlow. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 16, 17 pages. https://doi.org/10.1145/3411764.3445354

[66] Financial Times. 2021. Charts that work: FT visual vocabulary guide. https://www.ft.com/content/c7bb24c9-964d-479f-ba24-03a2b2df6e85. [last accessed: Aug 22, 2023].

[67] Turo I Uskali and Heikki Kuutti. 2015. Models and streams of data journalism. *The journal of media innovations* 2, 1 (2015), 77–88.

[68] Andreas Veglis and Charalampos Bratsas. 2017. Towards a taxonomy of data journalism. *Journal of Media Critiques* 3, 11 (2017), 109–121.

[69] Wibke Weber, Martin Engebretsen, and Helen Kennedy. 2018. Data stories: Rethinking journalistic storytelling in the context of data journalism. *Studies in communication sciences* 2018, 1 (2018), 191–206.

[70] Shangyuan Wu, Edson C Tandoc Jr, and Charles T Salmon. 2019. Journalism Reconfigured: Assessing Human–machine Relations and the Autonomous Power of Automation in News Production. *Journalism studies* 20, 10 (2019), 1440–1457.

[71] Leni Yang, Xian Xu, XingYu Lan, Ziyan Liu, Shunan Guo, Yang Shi, Huamin Qu, and Nan Cao. 2022. A Design Space for Applying the Freytag's Pyramid Structure to Data Stories. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2022), 922–932. https://doi.org/10.1109/TVCG.2021.3114774

[72] Ji Soo Yi, Youn ah Kang, John Stasko, and J.A. Jacko. 2007. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics* 13, 6 (2007), 1224–1231. https://doi.org/10.1109/TVCG.2007.70515

[73] Ogan M Yigitbasioglu and Oana Velcu. 2012. A review of dashboards in performance management: Implications for design and research. *International Journal of Accounting Information Systems* 13, 1 (2012), 41–59.

[74] Mary Lynn Young, Alfred Hermida, and Johanna Fulda. 2018. What makes for great data journalism? A content analysis of data journalism awards finalists 2012–2015. *Journalism practice* 12, 1 (2018), 115–135.

[75] Yixuan Zhang, Yifan Sun, Lace Padilla, Sumit Barua, Enrico Bertini, and Andrea G Parker. 2021. Mapping the Landscape of COVID-19 Crisis Visualizations. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 608, 23 pages. https://doi.org/10.1145/3411764.3445381

CATALOG OF DATA-DRIVEN NEWS ARTICLE SAMPLES

[G1] The Guardian. Elections 2022: live council results for England, Scotland and Wales. 2022. https://www.theguardian.com/politics/ng-interactive/2022/may/05/elections-2022-results-live-local-council-england-scotland-wales

[G2] The Guardian. Truss or Sunak? Who do bookies favour to be the next PM? . 2022. https://www.theguardian.com/politics/2022/jul/27/tory-leadership-latest-odds-tracker-bookies-next-pm-sunak-truss

[G3] The Guardian. How did your MP vote on the fracking motion? 2017. https://www.theguardian.com/environment/ng-interactive/2022/oct/19/how-did-your-mp-vote-on-the-fracking-motion

[G4] The Guardian. Covid in the UK cases, deaths and vaccinations – the latest numbers. 2022. https://www.theguardian.com/world/2022/jun/01/covid-19-uk-cases-deaths-vaccinations-latest-numbers-coronavirus-national-data

[G5] The Guardian. Liz Truss v Rishi Sunak a timeline of their roads to Tory final two. 2022. https://www.theguardian.com/politics/ng-interactive/2022/aug/15/liz-truss-v-rishi-sunak-a-timeline-of-their-roads-to-tory-final-two

[G6] The Guardian. The worldwide scale of the Uber files – in numbers. 2022. https://www.theguardian.com/news/2022/jul/15/the-worldwide-scale-of-the-uber-files-in-numbers

[G7] The Guardian. Europe's record summer of heat and fires – visualised. 2022. https://www.theguardian.com/environment/ng-interactive/2022/jul/26/how-europe-has-been-hit-by-record-fire-damage-and-temperatures

[G8] The Guardian. Boris Johnson's tumultuous three years as prime minister – in charts. 2022. https://www.theguardian.com/politics/2022/jul/07/boris-johnsons-tumultuous-three-years-as-prime-minister-in-charts

[G9] The Guardian. Are enough people vaccinated in time for England's 'freedom day'? 2021. https://www.theguardian.com/world/2021/jul/19/are-enough-people-vaccinated-in-time-for-englands-freedom-day

[G10] The Guardian. Covid vaccine figures lay bare global inequality as global target missed. 2021. https://www.theguardian.com/global-development/2022/jul/21/covid-vaccine-figures-lay-bare-global-inequality-as-global-target-missed

[G11] The Guardian. In numbers what is fuelling Britain's cost of living crisis. 2022. https://www.theguardian.com/business/2022/feb/03/in-numbers-britains-cost-of-living-crisis

[G12] The Guardian. Which countries are the most (and least) committed to reducing inequality. 2017. https://www.theguardian.com/inequality/datablog/2017/jul/17/which-countries-most-and-least-committed-to-reducing-inequality-oxfam-dfi

[G13] The Guardian. Why are some areas of the UK in drought A visual guide. 2022. https://www.theguardian.com/environment/2022/aug/19/why-areas-uk-drought-visual-guide

[G14] The Guardian. More than 25m drink from the worst US water systems, with Latinos most exposed. 2021. https://www.theguardian.com/us-news/2021/feb/26/worst-us-water-systems-latinos-most-exposed

[Ft1] Financial Times. Tory leadership election tracker: Rishi Sunak and Liz Truss remain in race to be PM. 2022. https://www.ft.com/content/5ed407bd-ac0d-4293-a0fe-c01dfa15d9d2

[Ft2] Financial Times. Covid-19 vaccine tracker: the global race to vaccinate. 2022. https://ig.ft.com/coronavirus-vaccine-tracker/?areas=gbr&areas=isr&areas=usa&areas=eue&areas=are&areas=chn&areas=chl&cumulative=1&doses=total&populationAdjusted=1

[Ft3] Financial Times. Biden vs Trump:live result 2020. 2020. https://ig.ft.com/us-election-2020/

[Ft4] Financial Times. High UK inflation hits sales, confidence and activity. 2022. https://www.ft.com/content/ec8c3bfb-5473-4776-b483-f0a531940c46

[Ft5] Financial Times. How the coronavirus pandemic has disrupted the US Democratic primary calendar. 2020. https://www.ft.com/content/f3bb0944-4437-11ea-abea-0c7a29cd66fe

[Ft6] Financial Times. UK general election poll tracker. 2019.https://www.ft.com/content/263615ca-d873-11e9-8f9b-77216ebe1f17

[Ft7] Financial Times. Climate graphic of the week Arctic warming four times faster than rest of the planet. 2022.https://www.ft.com/content/9bb32c6f-b3dc-46bd-944e-9e583407b866

[Ft8] Financial Times. Older population in England and Wales hits record high. 2022.https://www.ft.com/content/c8334a2a-cef9-4ed8-9470-5a4dbcfcd351

[Ft9] Financial Times. Carmakers' battery plans in peril as raw material costs soar. 2022.https://www.ft.com/content/b4002e49-07ce-41d8-9d3b-b6ed55af798c

[Ft10] Financial Times. Lab-grown meat maker Eat Just unable to capitalise on Malaysia chicken ban. 2022.https://www.ft.com/content/f986f084-9bb7-4f3a-a91b-1a5c783fc257

[Ft11] Financial Times. UK public finances on 'unsustainable path', says OBR. 2022.https://www.ft.com/content/ae63f4fb-184e-42b0-8d46-7a8a9be89cf4

[Ft12] Financial Times. First round of 2022 French election in charts. 2022.https://www.ft.com/content/2dd79519-c478-4660-886e-316c5cb553b3

[Ft13] Financial Times. Global inflation tracker see how your country compares on rising prices. 2023.https://www.ft.com/content/088d3368-bb8b-4ff3-9df7-a7680d4d81b2

[Ft14] Financial Times. Can the EU wean itself off Russian gas. 2022.https://ig.ft.com/europes-race-to-replace-russian-gas/

[Ft15] Financial Times. The Queen's 70-year reign in 10 charts. 2022.https://www.ft.com/content/4ef98f0a-47c2-47e9-b93a-7e39f1d2a4d7

[Ft16] Financial Times. Would carbon food labels change the way you shop? 2022.https://ig.ft.com/carbon-food-labelling/

[Ft17] Financial Times. Why 'the worker's market' is not delivering for Americans? 2022.https://www.ft.com/content/5a13b24b-a16c-429c-9128-d48dd183be5c

[Ft18] Financial Times. Where are all the workers? The US states and sectors with the tightest jobs markets. 2022.https://www.ft.com/content/9c5b5025-a995-494c-83a5-65ef04952a1a

[Ft19] Financial Times. Baby bust: economic stimulus helps births rebound from pandemic. 2022.https://www.ft.com/content/32436917-00b8-447d-8d6c-41f4be72b03f

[Ft20] Financial Times. The global stagflation shock of 2022: how bad could it get? 2022.https://www.ft.com/content/d490ef4e-3187-471e-84ff-9c065871a1a5

[B1] BBC News. Covid map: Coronavirus cases, deaths, vaccinations by country. 2022.https://www.bbc.co.uk/news/world-51235105

[B2] BBC News. What is the energy price cap and how high will bills go? 2023.https://www.bbc.com/news/business-58090533

[B3] BBC News. Brexit deal: How did my MP vote on the Withdrawal Agreement Bill? 2019.https://www.bbc.co.uk/news/uk-politics-50145265

[B4] BBC News. Coronavirus: What is the R number and how is it calculated? 2021.https://www.bbc.co.uk/news/health-52473523

[T1] The Times. World Cup 2022 Results, fixtures, stats and tables. 2022.https://www.thetimes.co.uk/sport/football/world-cup/fixtures-results

[T2] The Times. Don't Trust Your Gut by Seth Stephens-Davidowitz — can data make us happy? 2022.https://www.thetimes.co.uk/article/dont-trust-your-gut-by-seth-stephens-davidowitz-can-data-make-us-happy-kfjm9926l

[T3] The Times. Digital banks beat the high street on customer service. 2022.https://www.thetimes.co.uk/article/online-banks-outrank-high-street-rivals-for-customer-service-x6jwsxt5r

[T4] The Times. Hospital treatment is still a waiting game. 2022.https://www.thetimes.co.uk/article/hospital-treatment-is-still-a-waiting-game-pk5d85bfz

[T5] The Times. The five charts that explain the energy crisis. 2022.https://www.thetimes.co.uk/article/the-five-charts-that-explain-the-energy-crisis-lp8h3gcs6

[T6] The Times. Tory members are making the decision - but do they represent public opinion. 2022.https://www.thetimes.co.uk/article/tory-members-are-making-the-decision-but-do-they-represent-public-opinion-m80jnscqq

[T7] The Times. How much did the Platinum Jubilee cost and who is paying. 2022.https://www.thetimes.co.uk/article/how-much-did-the-platinum-jubilee-cost-and-who-is-paying-ww28ptqbf

[T8] The Times. What is the UK's true Covid death toll. 2022.https://www.thetimes.co.uk/article/what-is-the-uks-true-covid-death-toll-lzpb0xnqs
[E1] The Economist. Does Gen Z spend too much time on social media? 2022.https://www.economist.com/graphic-detail/2022/08/10/does-gen-z-spend-too-much-time-on-social-media
[E2] The Economist. Business. 2022.https://www.economist.com/the-world-this-week/2022/08/18/business
[E3] The Economist. Business. 2022.https://www.economist.com/the-world-this-week/2022/07/14/business
[E4] The Economist. Global living standards are moving in the wrong direction. 2022.https://www.economist.com/graphic-detail/2022/09/08/global-living-standards-are-moving-in-the-wrong-direction
[E5] The Economist. Elizabeth II was the longest-reigning monarch in British history. 2022.https://www.economist.com/graphic-detail/2022/09/08/elizabeth-ii-was-the-longest-reigning-monarch-in-british-history
[E6] The Economist. The world is almost back to pre-covid activity levels. 2022.https://www.economist.com/graphic-detail/2022/09/08/the-world-is-almost-back-to-pre-covid-activity-levels
[E7] The Economist. Another mass shooting in America. 2022.https://www.economist.com/graphic-detail/2022/07/05/another-mass-shooting-in-america
[E8] The Economist. Surging fossil-fuel emissions are ruining carbon dating. 2022.https://www.economist.com/graphic-detail/2022/08/12/surging-fossil-fuel-emissions-are-ruining-carbon-dating
[E9] The Economist. Better measurement would help reduce water consumption. 2022.https://www.economist.com/britain/2022/08/11/better-measurement-would-help-reduce-water-consumption
[E10] The Economist. The Big Mac index. 2022.https://www.economist.com/big-mac-index?fsrc=core-app-economist
[E11] The Economist. Friendship across class lines may boost social mobility and decrease poverty. 2022.https://www.economist.com/graphic-detail/2022/08/11/friendship-across-class-lines-may-boost-social-mobility-and-decrease-poverty
[E12] The Economist. Tracking covid-19 excess deaths across countries. 2022.https://www.economist.com/graphic-detail/coronavirus-excess-deaths-tracker
[E13] The Economist. What Spotify data show about the decline of English. 2022.https://www.economist.com/interactive/graphic-detail/2022/01/29/what-spotify-data-show-about-the-decline-of-english
[E13] The Economist. What Spotify data show about the decline of English. 2022.https://www.economist.com/interactive/graphic-detail/2022/01/29/what-spotify-data-show-about-the-decline-of-english
[E14] The Economist. There have been 7m-13m excess deaths worldwide during the pandemic. 2021.https://www.economist.com/briefing/2021/05/15/there-have-been-7m-13m-excess-deaths-worldwide-during-the-pandemic
[E15] The Economist. How we estimated the true death toll of the pandemic. 2021.https://www.economist.com/graphic-detail/2021/05/13/how-we-estimated-the-true-death-toll-of-the-pandemic
[E16] The Economist. Why it is too early to say the world economy is in recession. 2022.https://www.economist.com/finance-and-economics/2022/07/24/why-it-is-too-early-to-say-the-world-economy-is-in-recession
[NYT1] The New York Times. Coronavirus in the U.S.: Latest Map and Case Count. 2021.https://www.nytimes.com/interactive/2021/us/covid-cases.html
[NYT2] The New York Times. Tracking Dangerous Heat in the U.S.. 2022.https://www.nytimes.com/interactive/2022/us/heat-wave-map-tracker.html
[NYT3] The New York Times. Path to 218 Tracking the Remaining House Races. 2022.https://www.nytimes.com/interactive/2022/11/10/us/elections/results-house-seats-elections-congress.html?searchResultPosition=1
[NYT4] The New York Times. Live Election Results: Top Races to Watch. 2022.https://www.nytimes.com/interactive/2022/11/08/us/elections/results-key-races.html?action=click&pgtype=Article&state=default&module=election-results&context=election_recirc&region=NavBar
[NYT5] The New York Times. World Cup 2022: How Danmark can advance to the round of 16. 2022.https://www.nytimes.com/interactive/2022/upshot/denmark-world-cup-scores-standings.html?action=click&algo=identity&block=more_in_recirc&fellback=false&imp_id=39268242&impression_id=b0d7fac1-705c-11ed-8bf3-b7515042b3b5&index=1&pgtype=Article&pool=more_in_pools%2Ftheupshot&region=footer&req_id=822058229&surface=eos-more-in&variant=holdout_more-in
[NYT6] The New York Times. Coronavirus World Map: Tracking the Global Outbreak. 2021.https://www.nytimes.com/interactive/2021/world/covid-cases.html
[NYT7] The New York Times. What the BA.5 Subvariant Could Mean for the United States. 2022.https://www.nytimes.com/interactive/2022/07/07/us/ba5-covid-omicron-subvariant.html
[NYT8] The New York Times. Serena Williams: Charting a Career at the Top. 2022.https://www.nytimes.com/interactive/2022/08/09/sports/tennis/serena-williams-ranking-tennis-wins.html
[NYT9] The New York Times. U.S. Has Far Higher Covid Death Rate Than Other Wealthy Countries. 2022.https://www.nytimes.com/interactive/2022/02/01/science/covid-deaths-united-states.html
[NYT10] The New York Times. A Detailed Picture of What's in the Democrats' Climate and Health Bill. 2022.https://www.nytimes.com/interactive/2022/08/13/upshot/whats-in-the-democrats-climate-health-bill.html
[NYT11] The New York Times. How Nathan Chen Won Gold in Men's Figure Skating. 2022.https://www.nytimes.com/interactive/2022/02/10/sports/olympics/nathan-chen-jumps.html
[NYT12] The New York Times. The Dollar Is Extremely Strong, Pushing Down the World. 2022.https://www.nytimes.com/interactive/2022/07/16/business/strong-dollar.html
[NYT13] The New York Times. Where $5 Trillion in Pandemic Stimulus Money Went. 2022.https://www.nytimes.com/interactive/2022/03/11/us/how-covid-stimulus-money-was-spent.html
[NYT14] The New York Times. How the New Climate Bill Would Reduce Emissions. 2022.https://www.nytimes.com/interactive/2022/08/02/climate/manchin-deal-emissions-cuts.html
[NYT15] The New York Times. Vast New Study Shows a Key to Reducing Poverty: More Friendships Between Rich and Poor. 2022.https://www.nytimes.com/interactive/2022/08/01/upshot/rich-poor-friendships.html
[NYT16] The New York Times. The Exceptionally American Problem of Rising Roadway Deaths. 2022.httpswww.nytimes.com20221127upshotroad-deaths-pedestrians-cyclists.html
[NYT17] The New York Times. See Everything the White House Wanted, and Everything It Got. 2022.https://www.nytimes.com/interactive/2022/10/20/upshot/biden-budget-before-after-animation.html
[NYT18] The New York Times. 'The Cash Monster Was Insatiable' How Insurers Exploited Medicare for Billions. 2022.https://www.nytimes.com/2022/10/08/upshot/medicare-advantage-fraud-allegations.html