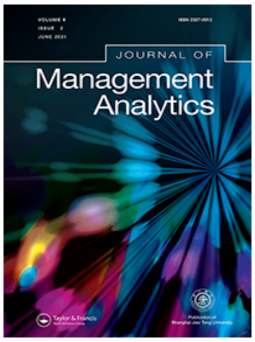


Journal of Management Analytics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjma20

Enhanced data narratives

Judd D. Bradbury & Rosanna E. Guadagno

To cite this article: Judd D. Bradbury & Rosanna E. Guadagno (2021) Enhanced data narratives, Journal of Management Analytics, 8:2, 171-194, DOI: 10.1080/23270012.2021.1886883

To link to this article: https://doi.org/10.1080/23270012.2021.1886883

Published online: 12 Mar 2021.

Submit your article to this journal

Article views: 26

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tjma20

Journal of Management Analytics, 2021
Vol. 8, No. 2, 171–194, https://doi.org/10.1080/23270012.2021.1886883

Enhanced data narratives

Judd D. Bradbury [a*] and Rosanna E. Guadagno [b]

[a] *Business Analytics Program, Jindal School of Management, University of Texas at Dallas, Richardson, TX, USA;* [b] *The Center for International Security and Cooperation, Stanford University, Stanford, CA, USA*

(*Received 17 December 2020; accepted 3 February 2021*)

Data narratives are an emerging form of communication that employs enhanced media for effective knowledge transfer of complex information. Researchers in the fields of data visualization and artificial intelligence have begun to pioneer new structures of communication to improve the efficiency of construction and the retention of information provided by the knowledge transfer experience. In this paper, we report the results of an empirical study conducted to compare the performance of various narrative communication techniques including frame based narrative visualization, documentary narrative visualization, computer generated text narratives and human generated text narratives. We assess the knowledge transfer performance for each of these data driven narrative structures. Across all conditions, an identical set of knowledge retention questions assessed participants' recall of details from their assigned narrative communication. Statistical analysis on group performance answering the knowledge retention questions revealed that some narrative communication techniques perform better with general audiences.

Keywords: analytic communication; data visualization; algorithmic storytelling; video analysis; enhanced data narratives

Introduction

Data narrative is a category of communication methods that utilize an underlying quantitative data set as content for a knowledge transfer experience using narrative structure. As technological advances increasingly facilitate access to "big data" (Manyika et al., 2011; McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2010), research teams are developing new methods for the construction of narrative communication using quantitative data sets. Researchers in the fields of data visualization and artificial intelligence are working to enhance historic traditions of human communication. The need for enhanced communication techniques has evolved in parallel to a rapidly expanding scope of data. Modern analytical methods have the capacity to process data sets with thousands of variables and millions of records (Elmqvist, Stasko, & Tsigas, 2008; Eriksson, Byrne, Johansson, Trygg, & Vikström, 2013; Keim, 2002; Shneiderman, 2008). How do you properly construct communication about four thousand different facets of a phenomenon that is one million records deep?

*Corresponding author. Email: judd.bradbury@utdallas.edu

© 2021 Antai College of Economics and Management, Shanghai Jiao Tong University

To effectively infer and communicate meaningful information from datasets of this size requires that millions of records be summarized and presented to accurately represent the insights gained from the data without significant loss of the meaning. Data narratives, in their emerging forms, are the early attempts at forming advanced methods of human communication with the capacity to summarize and encode these deep and complex data sets. Here we focus on evaluating enhanced data narratives with summarizations of large record data sets.

Natural language written by a human author is a basic building block of human communication. Humans have also created images and used them for visual communication for many millennia, using the unique ability to encode enormous amounts of information in a single picture. In the field of human- computer interaction (HCI), several methods of data narrative have emerged. This study focuses on three particular methods including *Visual Data Narratives, Documentary Data Narratives, and Computer Generated Text Data Narratives.* In this paper, we focus on these methods as part of a model for measuring the performance of enhanced data driven communications.

Visual Data Narratives summarize a quantitative data set through a process of graphically encoding data elements and presenting them in an overall visual frame (Hullman et al., 2013; Hullman & Diakopoulos, 2011; Segel & Heer, 2010; Ziemkiewicz & Kosara, 2008). Among others, visually encoded objects known as glyphs are presented in a visual view as representations of summarized variables and values from a quantitative data set (Borgo et al., 2013; Munzner, 2014; Ward, 2008). Information is communicated through a visual frame or a series of frames that engage the audience in a review of glyphs as representations of many data values. Documentary Data Narratives utilize a similar style of visually encoded data viewing with the enhanced structure of full motion video (Bradbury & Guadagno, 2020). Often the presentation of evidence in Documentary Data Narrative is delivered using verbal "voice-of-god" narration as a second channel of communication further improving the knowledge transfer experience. The enhanced capabilities provided by a moving video experience in Documentary Data Narrative opens up the possibility for exploration of animated data visualizations in a short film (Amini, Henry Riche, Lee, Hurter, & Irani, 2015; Kwon, Stoffel, Jäckle, Lee, & Keim, 2014; Robertson, Fernandez, Fisher, Lee, & Stasko, 2008; Yee, Fisher, Dhamija, & Hearst, 2001). Computer Generated Text Data Narratives summarize quantitative data sets with totals, ratios and values woven into a journalistic text-based story (CITO Research & Narrative Science, 2015). The story is composed using algorithmic text in place of the human author, constructing stories with typical phrases found in human communication.

Storytelling has been effective in human communication as expressed in the long and diverse history of teaching, education, and socialization of information using a narrative format (Green, 2004; Green & Brock, 2002; Oatley et. al., 2002). Utilization of narrative in business strategy allows for the exploration of characters, relationships, and cause and effect outcomes that simple bullet points cannot (Gershon & Page, 2001; Shaw, Brown, & Bromiley, 1998). At early age, we are taught to process information using the narrative arc of children's stories. We emulate these stories through the explanation of events that occur in our daily life. In essence we are programming each other to understand our lives using our characters, our relationships, and the retelling of outcomes from our decisions. Narrative, as a technique, helps us to place information in an appropriate context. Storytelling helps to mentally construct a world where information content lives, connecting that world to personal experience. In many ways, storytelling is the human operating system for exchanging complex information within a context. This begs the question: which of the extant enhanced approaches is the most effective technique for creating data narratives that produce a high degree of knowledge transfer? Below, we review the benefits of each data-driven approach.

Visual Data Narratives benefit from a significant ability to summarize and computationally encode complex data sets (Larkin & Simon, 1987). Quantitative data sets with a large scope can be engaged visually with navigation, layering, so-called small multiples, and multi-dimensional representation (Card, 1999; Munzner, 2009; Shneiderman, 1996; Yi, Kang, & Stasko, 2007). The challenge for Visual Data Narratives is in the scope of audience familiar with frame based presentation of data visualization. Frame based communication techniques have been used previously in comic books, storyboards, slide show presentations, new journalistic approaches, and business intelligence dashboards. While there is a significant amount of experience with frame based presentation of information, the experience of a general audience with data visualization presented in frame-based presentations with stepper navigation is significantly less than the experience of the same audience with communication in text narrative.

Documentary Data Narratives benefit from a large cross section of people experienced in watching films and from engaging two channels of information processing versus one (Mayer & Anderson, 1991; Mayer & Moreno, 1998; Mayer & Sims, 1994; Tversky, Morrison, & Betrancourt, 2002). Similar to Visual Data Narratives, Documentary Data Narratives are constructed from data visualization where audiences have significantly less experience interpreting the information. However, general audiences do have a great deal of experience watching documentary videos. When verbal narration (i.e. the "voice-of-god") is engaged, Documentary Data Narratives benefit even further from the long history of verbal storytelling containing a more universal reach of literacy.

Computer Generated Text Data Narratives benefit from looking and feeling similar to traditional *Human Generated Text Data Narratives.* Without clearly indicating the source of computer generated data narratives, research subjects may not be able to determine the difference between a data narrative generated by a human author versus one generated by a computer. Using the common communication technique of the written word, Computer Generated Text Data Narratives and Human Generated Text Data Narratives share a familiar knowledge transfer experience. Over a number of repetitive reading experiences, if the same construction template is used, audience members might develop commentary about boredom with the consistent pattern of the stories. This commentary is balanced by the positive feature of consistency of information communication in a series of stories. Based on literacy in general, Computer Generated Text Data Narratives have the ability to reach a very wide audience. Computer Generated Text Data Narratives have challenges when encoding a large scope of information in a concise way without the use of visual imagery. Text describes data well, but its nature may be more limited in its ability to summarize or computationally encode a larger scope of information.

Empirical research regarding the knowledge transfer performance of these methods for enhancing data narratives has yet to address which of these methods is superior. The purpose of this paper is to evaluate the knowledge transfer capability for each of these emerging types of enhanced data narratives. Specifically, we examine the knowledge transfer for each of these data narrative methods, alongside the classic example of text stories written by humans. We assess the recall performance for participants who viewed one of the four data narrative methods, communicating information about the common data set.

Specifically, we explore three primary research questions related to a general audience in this paper:

- Whether enhanced forms of data narrative transfer knowledge more effectively than traditional human generated text narratives?
- Are data narratives constructed with data visualizations more effective at knowledge transfer than text narratives?
- Are Documentary Data Narratives composed with video of data visualizations more effective at knowledge transfer than frame-based visualization data narratives?

Enhanced forms of data narrative

With rising expectations to create meaning from an expanding ocean of data, techniques for encoding and communicating information have begun to emerge. These techniques are designed to meet the challenge of synthesizing and summarizing large data sets into meaningful information for intellectual consumption. This study examines three taxonomic groups of data narrative construction defined here as Visual Data Narrative, Documentary Data Narrative, and Computer Generated Text Data Narrative. Each of three classes of data narrative uses communication techniques to represent a data set. Raw data sets are generated from observations about a phenomenon. The first two classes are defined as Visual Data Narrative and Documentary Data Narrative utilize data visualization to encode communication views representing the data set. The third class defined as computer-generated text data narrative summarizes data elements into quantitative numbers. The quantitative numbers are then mapped into common linguistic phrases using variables. Linguistic phrases are then composed into narrative paragraphs using standard subject matter templates (Narrative Science, 2015). The equivalent construction for a Human Generated Text Data Narrative has the author summarizing a data set into quantitative measures. The measures are then integrated into sentences and paragraphs necessary to create a narrative story.

Visual data narrative

Visual Data Narrative refers to a style of information presentation that constructs stories from frames of data visualization organized using a traditional narrative storytelling arc. Visual Data Narratives use a visual frame or view constructed from a data set as the building block for the overall construction of the data narrative. Data Visualizations in the Visual Data Narrative are typically constructed from an underlying data set using shapes, maps and other visual encoding to reveal a particular insight about a phenomenon represented by the data measurements. Views can also be constructed using data objects like text, pictures, and isotypes (small iconic pictures like gingerbread people) to help structure the overall narrative story (Haroz, Kosara, & Franconeri, 2015; Howson, 2014; SAP, 2015) Each frame of data visualization takes the shape of an event in a traditional story, linking cause and effect within time and space (Bordwell, Thompson, & Smith, 1997). Like narrative storytelling, Visual Data Narratives construct a series of data visualizations positioned as events, linked by cause and effect when appropriate, sequenced in time and located in space or geography.

One of the simplest methods for constructing a series of visualization events is defined in the Aristotelian tradition of a three-act play defined as a *beginning, middle,* and an *ending* (Aristotle, 1984). The beginning view creates a frame of reference for the audience. From the standpoint of spatial and temporal location, this first frame should passively help the audience locate their proximity to the phenomena represented in the Visual Data Narrative. The beginning might locate the audience within an industry, company, geography, virtual concept, and/or point in time. Rhetoric, for example using the Aristotelian modes of ethics, logic, or emotion can be engaged as a theme, bringing human meaning to the presentation. The introduction of rhetoric helps the audience identify with the narrative. How does this information relate to me? The middle view then uses a visualization view to highlight an issue or insight with a significant amount of tension. The more tension created by this middle view, the more effective the Visual Data Narrative. The ending presents a view of data that reconciles or summarizes the tension created during the middle view. Risk mitigation, alternate courses of action, and new opportunities are fertile approaches for reconciling the tension created in the middle view. Titles, headings, and annotations in each of the views help to provide narration of the insights defined in the data visualizations.

In cases where the content requires a more elaborate treatment of events, a structure similar to the classic Freytag Pyramid of *Setup, Rising Action, Climax, Falling Action,* and *Resolution* may provide a more interesting composition (Freytag, 1896). The rising action, climax, and falling action provide more intricate treatments of the presentation of tension in comparison to the simpler middle view of the Aristotelian model.

A more recent approach created by Neil Cohn is called *Visual Narrative Structure,* emulating the presentation approach in comic books (Cohn, 2013; McCloud, 1993). Visual narrative structure can be utilized to organize and sequence a series of cause and effect-linked data visualizations designed around unique styles of visual views. The visual narrative structure model sequences five styles of visual frames known as *establishers, initials, prolongations, peaks, and releases* (Cohn, 2013). *The Establisher* frame sets up the action in a scene. Characters are often introduced in the establisher frame. The *Initial* frame begins tension as a path of narrative arc. The *Initial* frame begins anticipation of the goal/peak of a path. The *Prolongation* view marks the middle of the narrative arc and is designed to expand anticipation in the story. This frame may be a somewhat passive or cyclical engagement of an on-going process. The *Peak* frame marks the high point of narrative tension as an event. The *Peak* often represents a critical event, or the goal of a path being achieved. The *Release* frame then relieves the tension created in the overall narrative arc.

For the present study, we constructed a Visual Data Narrative in the form of an interactive slide show with multiple frames of data visualization. The interactive slides are presented in a predefined sequence commonly referred to as a stepper in Tableau Storypoints. The categorization is derived from a navigation feature representing an arrow of time that allows the user to advance through data visualization frames, one step at a time. We selected this style of Visual Data Narrative as it is representative of a large cross-section of previous narrative visualization work used by the New York Times, The Guardian, and a number of other mass media publishers (New York Times, 2015; The guardian, 2015). The stepper approach allows the media publisher to check user attention through recorded clicks at particular steps. The interactive slide show with stepper navigation is also the designated style of storytelling represented in the *Storypoints* functionality in Tableau Software, as well as the *Stories* functionality in SAP Lumira Software (SAP, 2015; Tableau, 2015). The interactive slide show with stepper navigation is a leading archetype that is representative of Visual Data Narratives (Segel & Heer, 2010).

The Visual Data Narrative story presented in Figures 1–6 engage the more elaborate narrative structure models utilizing elements from the Freytag and Cohn approaches to visual narrative structure. The first frame in the story is presented as a *setup* or *establisher* to orient the audience to Intel as a one of the leading producers of microprocessors in the world using a simple line of text (Intel Corporation, 2015). The second frame is presented as an *initial* introducing the tension of profit decreasing in the previous quarterly financial results. A simple bar chart of four quarters was utilized to represent the underlying data measurements for quarterly profit. The third frame was presented as *rising action* increasing tension as revenue is also falling short of the previous year comparison. A bar chart referencing the quantitative measures of annual revenue are provided for comparison. The fourth frame is presented as the *peak* or *climax* of the story with earnings estimates increasing over a ninety-day period. A trend line chart is presented as the visualization containing three plotted points representing quantitative earnings measure estimates from investment analysts at three separate points in time. The fifth frame is presented as *falling action* with the quarterly and annual earnings measurements reported in alignment with analyst estimates. A bar chart of reported earnings measures for the current quarter and the year are presented beside earnings measures for the previous three quarters. The last frame is a *release* using simple text statements suggesting that analysts were optimistic about Intel with reported earnings reported consistent with expectations.

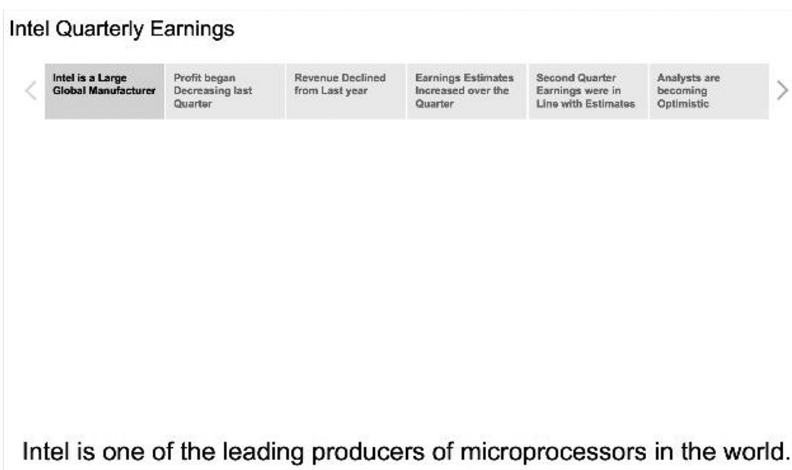


Figure 1. Visual data narrative of Intel earnings: setup-establisher.

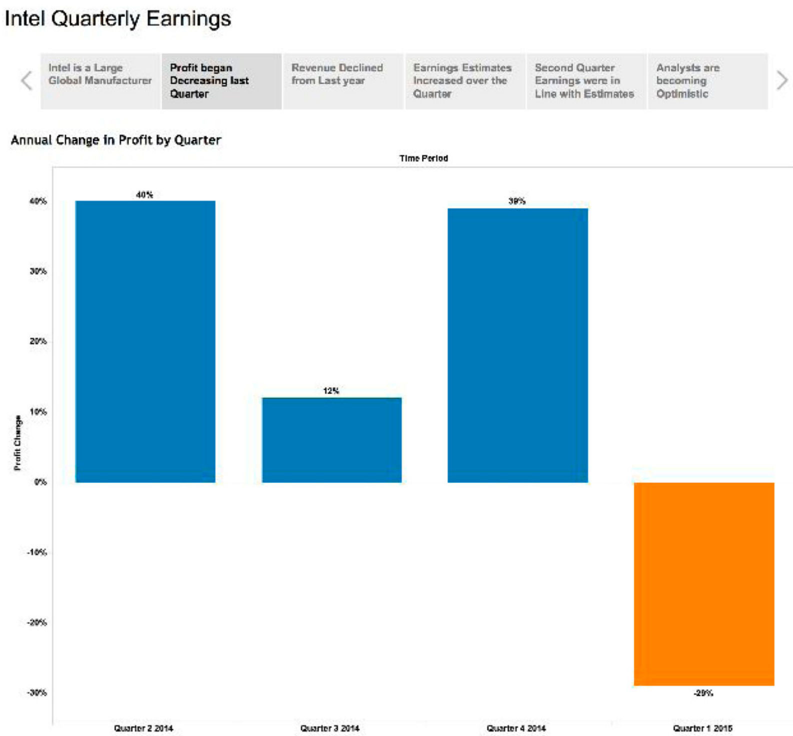


Figure 2. Visual data narrative of Intel earnings: initial.

Documentary data narrative

Documentary Narrative Visualization is the presentation of a data visualization experience using the techniques of documentary film (Bradbury & Guadagno, 2020). Engaging some of the techniques from Documentary film, the data visualization views are set in motion using video and voice-over narration to enhance knowledge transfer. Documentary Data Narratives can be defined as a video that presents indexical visual evidence using verbal rhetoric (Bradbury & Guadagno, 2020). The video is often structured in one of three modes of evidence presentation traditionally known as expository, participatory, and observational (Nichols, 2010). Expository videos structure evidence in a rhetorical frame, where the audience is transported to a world of evidence that closely resembles personal experience. Expository narrative visualization typically includes the features of indexical evidence, voice-of-god commentary, narrative storytelling, and rhetorical proofs. Participatory data narrative videos engage the visualization creator as a character in the narrative presentation of evidence. The visualization creator may also provide voice-of-god narration for the video. Observational videos present evidence in the most objective manner possible without voice-of-god narration from the visualization creator or any other source.

In our study, we constructed the Documentary Data Narrative example using the expository mode of documentary containing presentation of indexical evidence,

Figure 3. Visual data narrative of Intel earnings: rising action.

voice-of-god commentary, and narrative storytelling. Our indexical evidence was represented using the identical six visualization frames utilized in our construction of Visual Data Narrative with stepper navigation. The navigation of the visualization frames was set to motion using video instead of user guided stepping. The video was enhanced with voice-of-god narration engaging the text based insight annotations at the top of each visualization frame. Rhetorical enhancements were intentionally minimized to avoid injecting bias into our measurements.

Statements using the themes of ethics and emotion were avoided altogether. Careful attention was paid to include logical statements about insights that were consistent with the constructions of Visual Data Narrative and Computer Generated Text Data Narrative. See Figure 7 for the example of Documentary Data Narrative used in our study.

Computer generated text data narrative

Computer Generated Text Data Narrative is a form of communication that utilizes a computer algorithm to compose a text-based story. The composition process uses *Natural Language Generation (NLG),* to compose sentences and paragraphs from a quantitative data set by referring to a template of typical speech patterns that are common for an industry (Reiter & Dale, 1997). Natural Language Generation is a sub-field of artificial intelligence and computational linguistics that is concerned with the construction of computer systems that can produce understandable texts

Figure 4. Visual data narrative of Intel earnings: peak-climax.

in English or other natural languages from some underlying and non-linguistic representation of information (Reiter & Dale, 1997). The development of this type of automated construction of news stories is being pioneered by organizations like Narrative Science, Automated Insights, Yseop, and the Los Angeles Times (Automated Insights, 2015; Narrative Science, 2015; Slate, 2015; Yseop, 2015). As a leader in this sub-field, Narrative Science began with a business strategy of constructing journalistic stories for audiences of one that were cost prohibitive for human authors to write (Borgo et al., 2013). The Gamechanger website application constructs a sports news story for a little league team using an uploaded box score from the game (Allen, Templon, McNally, Birnbaum, & Hammond, 2010; Business Insider, 2014). This novel approach allows feature length stories to be constructed about little league baseball games without incurring the high labor cost of a professional journalist. In its second major initiative, Narrative Science created a partnership with Forbes magazine where the Narrative Science algorithm creates stories about earnings projections for publicly traded corporations that are published online a few days before the formal earnings announcements (The Atlantic, 2015). The earnings projections use a quantitative data set provided by Zacks Investment Research that is applied to a narrative template of financial sentence phrases using conditional if/then logical rules to construct a narrative text story about earnings that is then published in an online column in Forbes Magazine (Forbes, 2015). The algorithm driven investing column provides a low cost posting of a news story without the lower return on investment and tedious effort of a human author.

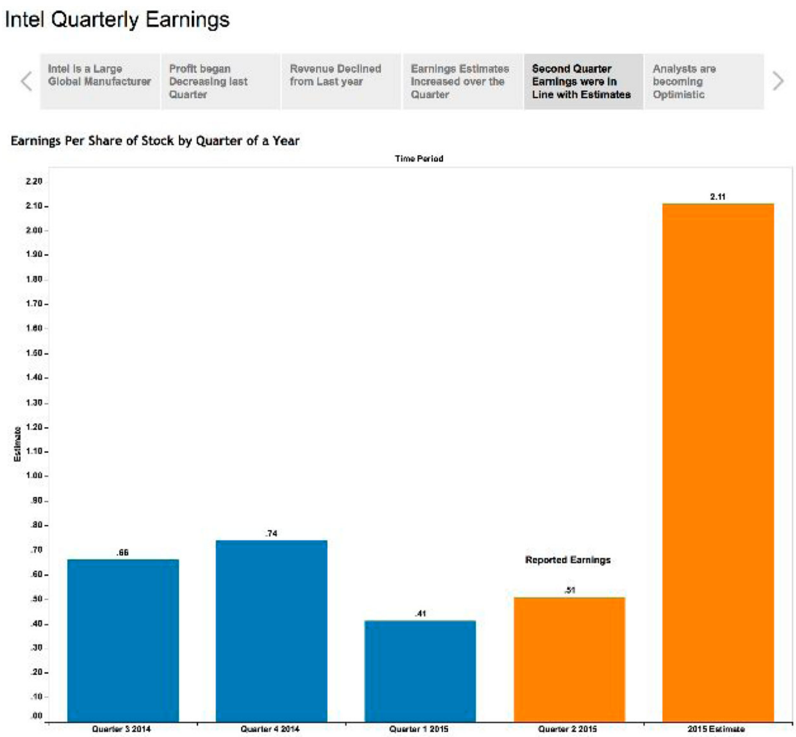


Figure 5. Visual data narrative of Intel earnings: falling action.

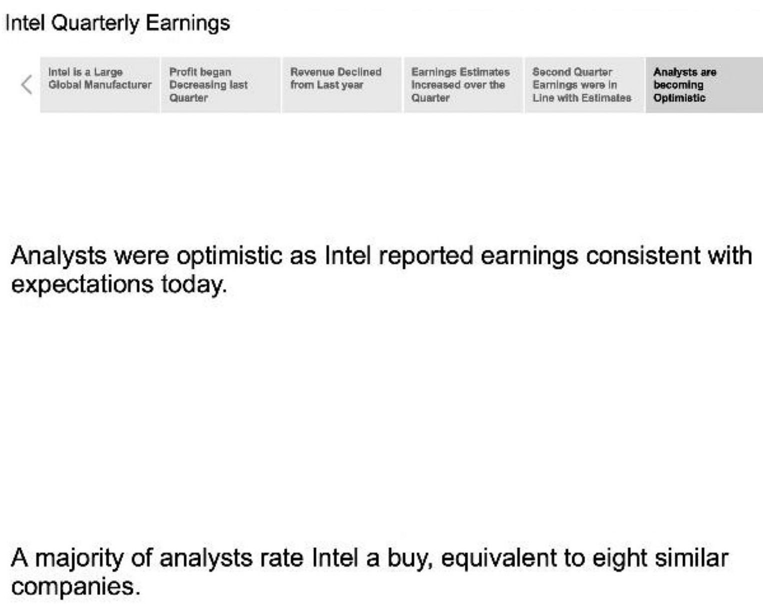


Figure 6. Visual data narrative of Intel earnings: release.

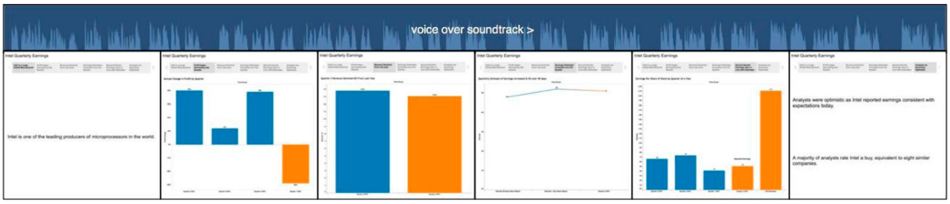


Figure 7. Intel earnings documentary data narrative with voice-of-god narration.

In our study, we utilized the Computer Generated Text Data Narrative composed by the Narrative Science algorithm that ran on July 13, 2015 in the electronic version of Forbes Magazine (Forbes, 2015). The text of the narrative was provided to our research participants in a single view as illustrated in Figure 8.

Forbes / Investing

JUL 13, 2015 @ 12:00 PM 464 VIEWS

Intel Profit Expected to Slip

Narrative Science, PARTNER

Narrative Science

Narrative Science

FOLLOW ON FORBES (15)

FULL BIO >

Opinions expressed by Forbes Contributors are their own.

Despite an expected dip in profit, analysts are generally optimistic about **Intel** as it prepares to reports its second-quarter earnings on Wednesday, July 15, 2015. The consensus earnings per share estimate is 51 cents per share.The consensus estimate has dipped over the past month, from 52 cents, but it's still up from the consensus estimate of 48 cents three months ago. For the fiscal year, analysts are expecting earnings of $2.11 per share. Analysts look for revenue to decrease 6% year-over-year to $13.07 billion for the quarter, after being $13.83 billion a year ago. For the year, revenue is projected to come in at $54.91 billion.A year-over-year revenue decrease in the first quarter snapped a streak of three consecutive quarters of revenue increases.The decline in profit in the first quarter followed three straight quarters of year-over-year profit growth. Profit dropped 29% year-over-year in the most recent quarter. Going further back, the figure rose 39% in the fourth quarter, 12% in the third quarter and 40% in the second quarter.The majority of analysts (59%) rate Intel as a buy. This compares favorably to the analyst ratings of eight similar companies, which average 54% buys.Intel develops advanced integrated digital technology products for industries such as computing and communications. Texas Instruments, also in the semiconductors industry, will report earnings on Wednesday, July 22, 2015. Analysts are expecting earnings of $0.65 per share for Texas Instruments, up 5% from last year's earnings of $0.62 per share. Other companies in the semiconductors industry with upcoming earnings release dates include: ARM Holdings plc and NVIDIA.*Earnings estimates provided by Zacks.Narrative Science, through its proprietary artificial intelligence platform, transforms data into stories and insights.*

Figure 8. Intel earnings computer generated text data narrative.

Human generated text data narrative

A human authored text data narrative was constructed for our study to provide a baseline of comparison for the enhanced data narrative communication conditions. Journalistic stories about corporate earnings are standard fare for financial news media services. The earnings announcement column is a typical assignment for new journalists in the financial industry. Our objective in engaging a Human Generated Text Data Narrative was the creation of an expert human authored story on corporate earnings using phrases typically used by human authors, edited to mirror the same quantitative measures presented in the enhanced data narrative presentations. The Human Generated Text Data Narrative for our study was constructed using sentences from two earnings articles published by the financial media websites of Zacks Investment Service and Seeking Alpha (Zacks, 2015).

Human communication proxy

The Mathematical Theory of Communication, presents a model of communication where transmitters send messages across communication channels to receivers (Shannon & Weaver, 1951). In the case of in person human communication, the transmitter is the speaker, the verbal voice over the air is the communication channel, and human listening is the receiver. Communication channels use presentational media including voice, spoken words, facial expressions, gestures; and representational media in the forms of books, text, paintings, and images; transmitted through technological media of telecommunications, internet, radio, and television (Fiske, 2010). The combinations of transmitters, receivers, and communication channels provide unique approaches to human communication.

The most common form of communication is found with one person interacting with another person during a face-to-face verbal exchange. Alternative approaches utilize technological media substitutes for various communication elements of the sender or receiver . Our findings in this paper speak to the concept of a Human Communication Proxy that works to emulate aspects of a one on one conversation between two people. Cognitively, the face-to-face human communication approach appears to function as a normative model of audience preference (Barkhi & Pirkul, 1999; Daft & Lengel, 1986; Daft, Lengel, & Trevino, 1987; Dennis, Fuller, & Valacich, 2008; Denstadli, Julsrud, & Hjorthol, 2012; Kock, 2004; Kydd & Ferry, 1991; Lee, 2010).

Our results suggest that general audience preferences for enhancements to data driven storytelling are being measured against the normative face-to-face human communication model. The HCP substitutes elements of presence, verbal narration, story structure, and representational media transmitted to the receiver. When the HCP is engaged, audience preference follows how rich or realistic the representation is evocative of the normative face-to-face sender/receiver experience. The closer the emulation is to a conversation with another person, the more attractive the communication approach is to a general audience. Our findings support media naturalness theory and the psychobiological model predicting that lower naturalness in the narrative experience translates to higher cognitive effort in consuming a narrative (Kock, 2004). Thus, the empirical results presented in this paper suggest that general audiences express a preference for communication models where they have more experience using them (Figure 9).

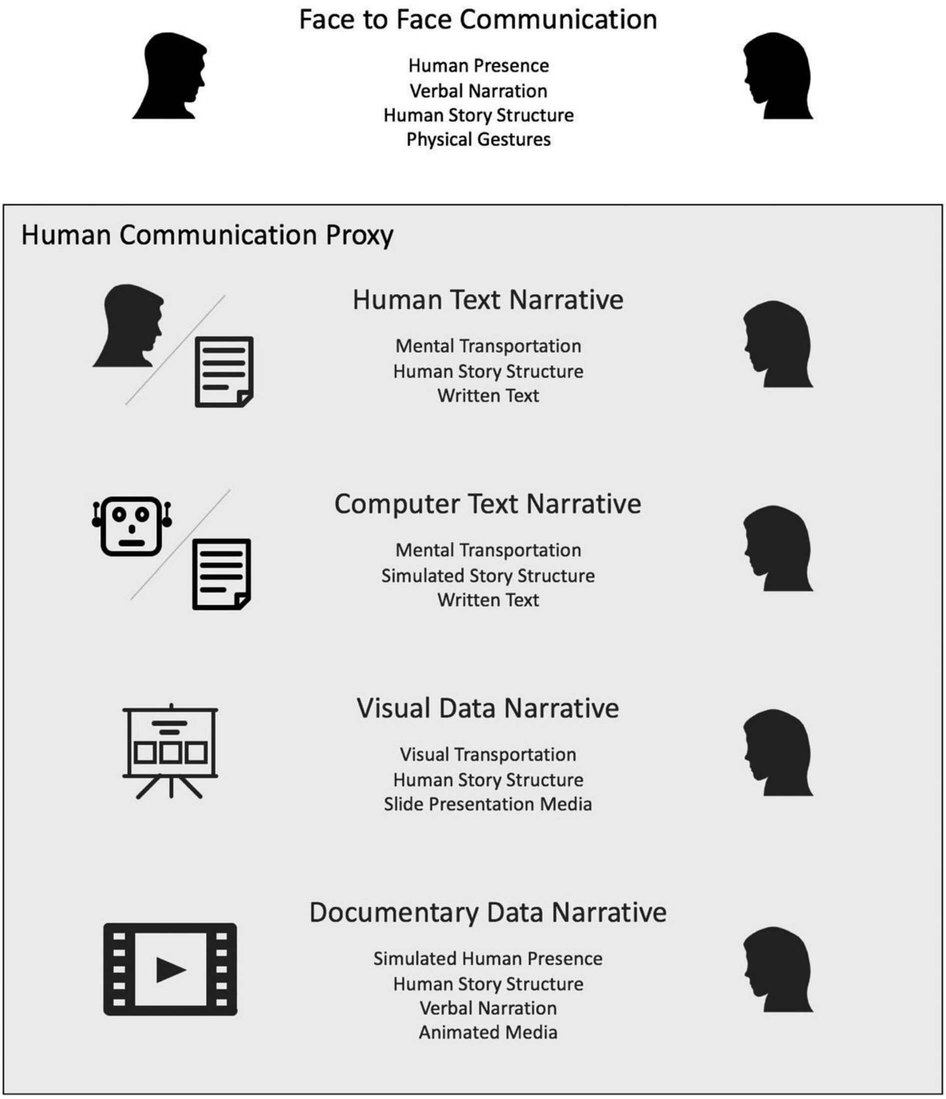


Figure 9. Human communication proxy.

Enhanced Data Narratives using Computer Generated Text utilize a communication model with an algorithmic author that crafts a text based representation of a story. The text based story is published or provided in some manner by an algorithmic enhanced sender. The published text media functions as the communication channel. The human receiver then consumes the text through a process of a human reading it, using any of several media including newsprint, computer screen, or smart device. The HCP in this model is found in the algorithmic story structure applied in the creation of the computer generated text. The written text also functions as a substitute for the in person narrative experience provided by a sender. As a surrogate for the in person narrative experience, the text narrative transports the reader to a narrative world (Gerrig, 1993; Green & Brock, 2002).

The enhancement of computer automated authoring is often not visible to the reader. With an optimized natural language generation algorithm, the experience of the receiver reading text may be indistinguishable from a similar experience reading a human authored equivalent. The only distinguishable difference for the reader might surface in the form of a predictable sentence pattern after reading several narratives written by the same algorithm. This phenomenon is also known to exist with formulaic human authors.

Visual Data Narratives utilize a communication model where a human author constructs a sequence of data visualizations in the form of a navigated stepper slide show. Once published, the author becomes the sender. The step-based views, visualizations, and rhetoric become the communication channel. The receiver interacts with the Visual Data Narrative to decode and consume the intended messages. The stepper slide show becomes the HCP standing in for the in person narrative experience. The standalone nature of the Visual Data Narrative makes it absent any form of verbal narration. This leaves all the communication to be performed through a visual only experience. The audience may be transported to a narrative world, but that world is silent, or limited to the narration provided from titles and annotations in the visual frames.

Documentary Data Narratives utilize a communication model where a visual data narrative is recorded and transformed into a data narrative video. Voice over narration is combined with a predetermined step through sequence of data visualization views. In most constructions, the creator becomes a character or narrator in the video. The creator publishes the recorded video that functions as the sender. The audio visual experience of the video playing is functioning as the communication channel. The receiver plays the video, consuming the communication by watching. The combination of voice over narration, visualizations, and in some cases visible human actors, altogether represent a Human Communication Proxy. The video engages visualization, titles and annotations in the visual channel of communication. Voice over narrative acts as a substitute for the in-person narrative experience, engaging the audio channel of communication. In concert, the video provides a higher fidelity communication experience that emulates the pacing and structure of a narrative conversation between two people.

Materials and methods

Our primary motivation in this study was to determine how effective each of the three emerging forms of data narrative is at transferring knowledge to a general audience. To accomplish this objective, we measured the knowledge transfer capability of the three major forms of data narrative using the most typical examples of each construction technique.

Design of the study

Using an identical data set, four narrative stories were constructed using each of the distinct communication techniques for narrative construction. The common data set utilized for each story construction was sourced from the financial data set for the Intel Corporation published by Zacks Investment Service (Zacks, 2015). Careful attention was made to ensure that the specific data elements used in each story were represented identically regardless of the type of narrative construction. The factual content and statements were the same in each of the data narrative constructions, providing focus on the construction techniques as dependent variable conditions for measurement.

In the first version, an example of Visual Data Narrative using Storypoints functionality in Tableau was created. The story building process constructs narratives using a stepper type function to advance from one data visualization frame to the next visualization in a sequence. The Visual Data Narrative was crafted with four data visualization frames of bar charts and a line chart. A blue and orange color scheme was utilized encoding blue objects as historical data and orange objects as current data. An introduction frame with a defining text sentence about Intel was sequenced in front of the four visualization frames, followed by a summarization frame with two text statements about the optimism and buy rating of financial analysts. A summarizing statement of insight was defined above each visualization frame in a light grey box, functioning as a sequential navigation button. A title was provided for each frame summarizing the content with the assistance of standard axis labelling and details where appropriate.

In the second version of our story, the Documentary Data Narrative used was constructed as a full motion video walk through of the same data visualization frames presented in the example of Visual Data Narrative. The example includes human voice over narrative describing the dynamic changes to the data represented in the data visualization.

In the third version, our source for the Computer Generated Text Data Narrative is an investing column projecting earnings for the Intel Corporation (Forbes, 2015). The earnings projection story was published a few days before the actual earnings were released. The article was written by an algorithm at Narrative Science and published online in Forbes Magazine (Forbes, 2015). The headline and text of this data narrative was presented to research subjects in the same manner that it is presented in the investing column at Forbes Magazine.

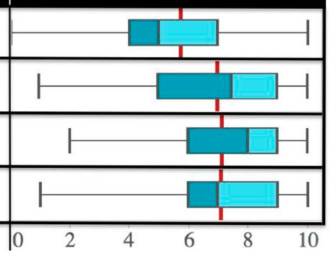
In the fourth version, the Human Text Data Narrative uses natural language written by a human author sourced from Zacks Investment Service from an article published for the equivalent Intel second quarter earnings announcement (Zacks, 2015). In cases where text sentences equivalent to the Computer Generated Text Data Narrative were unavailable in the Zacks article, sentences were sourced from an article published by Seeking Alpha for the Intel second quarter earnings announcement (Seeking Alpha, 2015). Our objective was the construction of a text narrative example containing identical content written by a human that was aligned as much as possible with the Algorithm Text Data Narrative. Sentences for the Human Text Data Narrative were incorporated from their article sources in their entirety. Sentences in the Human Text Data Narrative were edited to ensure the exact same quantitative data elements were presented.

The four types of data narrative communications along with descriptions of their unique features appear in Table 1. Visual Data Narrative creates the test condition listed first in the table with sequential frame based visualization construction. Documentary Data Narrative makes up the next test condition listed second in the table with full motion video construction. Computer Generated Text Data Narrative makes up the test condition listed third in the table with Natural Language Generated text construction. Human Generated Text Data Narrative is the fourth test condition listed in the table with human authored sentence construction.

Table 1. Recall performance of data narrative tests.

Recall Performance of Data Narrative Tests

Type of Data Narrative	N	Mean	Standard Deviation	Quiz Score
Visual Data Narrative	107	5.74*	2.26	
Documentary Data Narrative	130	6.98	2.32	
Computer Generated Text Data Narrative	134	7.11	2.17	
Human Generated Text Data Narrative	141	7.08	2.24	



Each of the four examples of data narrative were presented to research subjects in an equivalent manner with the narrative treatment residing within an equal sized 1024 by 768-pixel frame for viewing content. Instructions for reviewing each type of data narrative were provided at the top left of the viewing frame as a headline in bold. Study participants were randomly assigned to view one of the four data narrative constructions described in Table 1 and then asked to answer the same set of ten questions. Participants were allowed to engage the content at their own pace, with all participants spending a roughly equivalent amount of time on their narrative treatment. Participants were prevented from re- reviewing the data narrative content while answering the questions. Participants assigned to review the same data narrative construction were defined as a group of subjects representing that test condition.

Using the four groups of subjects, we applied a quantitative approach to examine the significance of differences in the average quiz score performance for each of the research subject groups.

Recall of results

A common data set was utilized to support each of the four types of data narrative content. The exact same quantitative measurements were defined in each type of data narrative. Where possible, sentences and headlines used the same size and font across all of the data narrative types, paying respect to the original intent of the type of data narrative construction. The factual statements used to construct the stimulus materials for each data narrative is described below:

- Intel is a large global manufacturer of microprocessors.
- Three quarters of increasing profits followed by the most recent quarter of profit declining by 29%.
- Revenue decreased this quarter by 6% to 13.07 billion dollars compared to 13.83 billion dollars in the same quarter of 2014.
- Consensus earnings estimates were .48 per share ninety days before the report, .52 per share one day before the report, and .51 per share the day of the report.
- Earnings per share of .51 were expected for the quarter, with analysts expecting $2.11 in earnings for the fiscal year.

- Analysts are becoming more optimistic as Intel reported earnings consistent with analyst expectations today.
- Most of the analyst's rate Intel a buy, equivalent to eight similar companies.

To assess recall, participants were asked to respond to 3 multiple choice questions about the trend of a measure, 6 multiple choice questions requesting recall of a specific quantitative measure, and 1 multiple choice question regarding the trend of analysts sentiment. The questions related to a trend were defined as:

- What happened to profit last quarter?
- What happened to revenue last quarter?
- What was the trend of earnings estimates?

All trend questions provided three randomly sequenced answer choices for the possible trend directions of increased, decreased, or stayed the same. The questions requiring quantitative measure recall were specified as:

- What was the amount of change to profit?
- What was the amount of reported revenue?
- What was the earnings estimate 90 days before the quarterly report?
- What was the earnings estimate 1 day before the quarterly report?
- What were the earnings reported in the second quarter of 2015?
- What were the earnings estimated to be for the entire year of 2015?

Four answer choices were presented randomly in a vertical sequence for each quantitative measure question. An example of the randomly sequenced quantitative measure choices for the question regarding earnings estimates in the year 2015 is found below:

- 3.16
- .89
- 1.84
- 2.11

The final question regarding the trend of analysts' sentiment was stated as?

- Are analysts becoming more optimistic or less optimistic about Intel?

Three randomly sequenced answer choices were provided for analysts' sentiment as more optimistic, less optimistic, and about the same. The ten questions were presented in the same sequence of profit questions, revenue questions, earnings questions, and finally the sentiment question.

Method

Our study was conducted to assess the knowledge transfer capabilities of emerging forms of data narrative. Data narrative constructions involving Visual Data Narrative, Documentary Data Narrative, Computer Generated Text Data Narrative, and

Human Text Data Narrative were presented as data driven communication for participants. Each participant was presented with one of the four data narrative constructions followed by a ten-item survey of questions designed to assess their recall of the information presented in the data narrative. Prior to reviewing the data narratives, participants were asked a series of preliminary questions to determine their interest in financial investments, their knowledge in financial matters, and their opinion of the products and reputation of the Intel Corporation. The survey concluded with a short set of demographic questions regarding standard indicators of age, gender, income, industry, level of education, and occupation.

Participants and procedure

Participants for the study were recruited using the Amazon Mechanical Turk platform with worker requirements defining US Citizens with at least a 98% approval rate for their hits (work items), that have completed more than 100 hits (Heer & Bostock, 2010). Each participant received $0.50 for engaging the study. Upon completion, participants were prevented from performing the study more than once using their Mechanical Turk ID. Participants began the survey by clicking on a link and providing their informed consent. A tracking feature was used to ensure participants were reviewing the narrative content they were assigned. We identified 5 participants that diligently reviewed the content of their assigned data narrative and completed all survey questions, except those regarding demographics. We respect the decision of these participants to avoid providing their demographic information. We included their responses for the questions not related to demographics in the results of the study. Participants that did not review the content of their assigned data narrative or did not answer the survey questions for knowledge recall were excluded from the study.

A group of 512 adult participants (261 men, 246 women, 5 unreported) from the United States were included in the study. They ranged in age from 18 to 74 (M = 35.84, SD = 11.35). Participants reported the following education levels: 43.8% attained a 4-year College Degree, 18.8% completed some college, 11.9% attained a 2-year College Degree, 10.9% attained a master's degree, 8.8% completed High School, 2.1% attained a Professional Degree, and 2.1% attained a Doctoral Degree. The three occupational categories of Legal, Computer/Mathematical, and Sales accounted for 38.1% of the participants in aggregate. More than 61% of the participants were drawn from the following industries: Legal (13.9%), Computer/Mathematical (13.1%), Sales (11.1%), Office or Administration (7.8%), Arts Design & Entertainment (7.8%), and Management (7.4%). Reported personal income was under $25,000 for 20.3% of the participants, with 68.5% classified in the $25,000 to $100,000 range, and 11.2% classified greater than $100,000.

We assessed participant responses regarding their interest in financial information, their knowledge of financial issues, as well as their attitudes about the Intel Corporation. For the participants surveyed, 92.6% exhibited at least some interest in financial information, with 81.8% having some knowledge of financial issues. Less than 3% of the participants had a negative bias about the Intel Corporation and its products.

Participants were randomly assigned to view one of the four data narrative test conditions (Group 1: Frame Based Visualization Story; Group 2: Video Visualization

Story; Group 3: Computer Authored Text Story; Group 4: Human Authored Text Story). Answers to the 10 questions assessing recall/understanding formed the scoring baseline for comparison of the groups.

Results

A one-way between-subjects analysis of variance (ANOVA) was conducted to explore the impact of data narrative construction on knowledge transfer, as measured by the survey questions on recall for the factual content of the data narratives. Participants were divided into four groups according to the type of data narrative they reviewed (Group 1: Visual Data Narrative; Group 2: Documentary Data Narrative; Group 3: Computer Generated Text Data Narrative; Group 4: Human Text Data Narrative). There was a significant main effect of data narrative type on recall scores for the four groups: $F(3, 508) = 9.84$, $p < .001$, $\eta_p^2 = .05$. The results of the post-hoc analysis is presented in Table 1.

Post-hocs using the Tukey Honest Significant Difference (HSD), highlighted by an asterisk in Table 1, indicated that participants' mean recall scores for the Visual Data Narrative type of construction performed significantly lower with our test audience than the other three types of data narrative construction. Mean recall scores in Groups 2, Group 3 and Group 4 did not differ significantly from each other.

Bonferroni corrections were applied to the p-values for the relevant comparison groups. The null hypothesis of a Computer Generated Text Data Narrative being as effective at knowledge transfer as a Human Generated Text Data Narrative was affirmed with a value of $p = .9992971$, with a correction of (1.000000). Visual Data Narratives underperformed Human Text Data Narratives with a value of $p = .0000246$, with a correction of (.000025). Visual Data Narratives also underperformed Computer Generated Text Data Narratives with a value of $p = .0000181$, with a correction of (.000018). Documentary Data Narratives significantly outperformed Visual Data Narratives with a value of $p = .0001470$, with a correction of (.00015).

Discussion

In this paper, we began with the purpose of assessing various forms of data narrative that have emerged over the last several years. Data Visualization, Narrative Visualization, Data Stories, and Natural Language Generated Text Stories have become very popular in the field of management. A great deal of time and energy has been devoted to these increasingly used methods. Our analysis was designed to apply an empirical approach to determine if enhanced data narrative communication techniques worked as well as traditional text narratives, and if they did, which ones worked better, and why?

Our first research question was an inquiry about the ability of enhanced forms of data narrative transferring knowledge better than traditional human generated text. In the case of Computer Generated Text Data Narratives, our results were consistent with our expectation. We expected participants to understand and recall the information presented in the Computer Generated Text Data Narrative at least as well as the Human Generated Text Data Narrative. After reviewing the Computer Generated Text Data Narrative, participants scored as well on the ten survey questions as the participants that had reviewed a Human Generated Text Data Narrative. Our results suggest that our example of Computer Generated Text Data Narrative, as an enhanced form of data narrative, did not perform better or worse than our example of a data narrative written by a human. In terms of utility, a Computer Generated Text Data Narrative is not necessarily designed with the objective of improving knowledge transfer compared to a human authored text data narratives. The objective of Computer Generated Text Data Narrative is providing a communication medium that is as effective at knowledge transfer as a human authored text narrative. The benefit of using Computer Generated Text Data Narratives is located in the cost savings of not hiring a human to write the data narrative. Given this specific objective, Computer Generated Text Data Narratives performed as well as Human Generated Text Data Narratives in our study.

In the case of Visual Data Narratives, we expected Visual Data Narratives to exhibit a superior capability of knowledge transfer. The results of our study suggest that the Visual Data Narrative we created underperformed their text-based counterparts to a significant degree. Participants that reviewed the example Visual Data Narrative content scored much lower on the ten survey questions requesting recall of information presented in the data narrative. Participants that reviewed the example Visual Data Narrative scored significantly lower on average than the participants that reviewed the example Human Generated Text Data Narrative and Computer Generated Text Data Narrative. The results of our study suggest that for the chosen examples we created, the model of Visual Data Narrative did not transfer knowledge as effectively as the examples of text based data narratives, when reviewed by a general audience that may lack advanced visual data literacy.

Our second research question asks if data narratives constructed with data visualizations are more effective at knowledge transfer than text narratives. Our study had mixed results on this question. When our example of frame based Visual Data Narratives was used without the enhancements of documentary film, the results of our study suggested that the Visual Data Narrative did not perform better at knowledge transfer than the text data narrative counterparts. On the other hand, our results suggest that when our example of Visual Data Narrative is constructed using video presentation, accompanied with verbal voice-of-god narration, the recall performance of participants improves to the same level as our examples of text data narratives generated by a human or computer.

Our third research question explored if Documentary Data Narratives are more effective at knowledge transfer than Visual Data Narratives. Our results suggest that our examples of Documentary Data Narrative communication techniques outperformed our examples of Visual Data Narrative communication techniques to a significant degree. Participants that were asked to review the Documentary Data Narrative example scored significantly higher on the ten survey questions requesting recall from the information presented, when compared to the group of participants that reviewed the same content presented in a stepper click based Visual Data Narrative approach.

Our original aspiration in developing this study was to provide some empirical evidence in favor of enhanced techniques for information presentation. The results of our study suggest that enhanced techniques for information presentation do play a productive role in the presentation of complex information for a broad audience. There is a large body of expert knowledge in the academic community that suggests advanced technologies and techniques like Natural Language Generation, Narrative Visualization, and Documentary Narrative Visualization have enhanced abilities to encode and represent large and/or complex data sets better than traditional communication techniques like human authored natural language text. The challenge of these enhanced techniques may reside in the literacy, or lack of knowledge building experiences, of general audiences regarding these new forms enhanced data narrative communication (Carlson & Zmud, 1999; Lengler, 2006). General audiences may struggle with some of the communication techniques that are less familiar to them. Natural language text generated by a computer algorithm shares the most fundamental and common approach for human communication. The enhancement provided by Natural Language Generation is in the production of information content. Once that content is produced, it is available in the least common denominator, or most universal communication technique of text. If our example is instructive, Visual Data Narratives may have more challenges with broader audiences, that more often than not, have no formal training in data visualization. This poses unique challenges when Visual Data Narratives are utilized for general audiences. The challenge of visual literacy in general audiences can be mitigated when Visual Data Narratives are enhanced with communication techniques that may be more familiar to the audience. Specifically, adding voice-of-god verbal narration, and video presentation seems to benefit from audience's experience with film, the comfort of a human voice, and a second cognitive channel of communication.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Judd D. Bradbury http://orcid.org/0000-0003-3664-2315

References

Allen, N. D., Templon, J. R., McNally, P. S., Birnbaum, L., & Hammond, K. (2010, November). Statsmonkey: A data-driven sports narrative writer. In *2010 AAAI Fall Symposium Series*.

Amini, F., Henry Riche, N., Lee, B., Hurter, C., & Irani, P. (2015, April). Understanding data videos: Looking at narrative visualization through the cinematography lens. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 1459–1468). ACM.

Aristotle, R. W. (1984). *Rhetoric*. New York: Random House.

Automated Insights. (2015). Introducing wordsmith. Automated insights website. Retrieved March 1, 2015. http://automatedinsights.com/wordsmith/

Barkhi, R., & Pirkul, H. (1999). An experimental analysis of face to face versus computer mediated communication channels. *Group Decision and Negotiation*, *8*(4), 325–347.

Bordwell, D., Thompson, K., & Smith, J. (1997). *Film art: An introduction*. New York: McGraw-Hill.

Borgo, R., Kehrer, J., Chung, D. H., Maguire, E., Laramee, R. S., Hauser, H., & Chen, M. (2013, May). Glyph-based visualization: Foundations, design guidelines, techniques and applications. In *Eurographics (STARs)* (pp. 39–63).

Bradbury, J. D., & Guadagno, R. E. (2020). Documentary narrative visualization: Features and modes of documentary film in narrative visualization. *Information Visualization*, *19*(4), 339–352.
Business Insider. (2014). If you don't think robots can replace journalists, check out this article written by a computer. Business Insider Website. http://www.businessinsider.com/narrative-science-quill-gamechanger-2014-7
Card, M. (1999). *Readings in information visualization: Using vision to think*. San Francisco: Morgan Kaufmann.
Carlson, J. R., & Zmud, R. W. (1999). Channel expansion theory and the experiential nature of media richness perceptions. *Academy of Management Journal*, *42*(2), 153–170.
CITO Research & Narrative Science. (2015). The automated analyst: transforming data into stories with advanced natural language generation. Narrative Science Whitepaper. Retrieved March 1, 2015. http://resources.narrativescience.com/h/i/114675300-white-paper-the-automated-analyst
Cohn, N. (2013). Visual narrative structure. *Cognitive Science*, *37*(3), 413–452.
Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, *32*(5), 554–571.
Daft, R. L., Lengel, R. H., & Trevino, L. K. (1987). Message equivocality, media selection, and manager performance: Implications for information systems. *MIS Quarterly*, *11*(3), 355–366.
Dennis, A. R., Fuller, R. M., & Valacich, J. S. (2008). Media, tasks, and communication processes: A theory of media synchronicity. *MIS Quarterly*, *32*(3), 575–600.
Denstadli, J. M., Julsrud, T. E., & Hjorthol, R. J. (2012). Videoconferencing as a mode of communication: A comparative study of the use of videoconferencing and face-to-face meetings. *Journal of Business and Technical Communication*, *26*(1), 65–91.
Elmqvist, N., Stasko, J., & Tsigas, P. (2008). Datameadow: A visual canvas for analysis of large-scale multivariate data. *Information Visualization*, *7*(1), 18–33.
Eriksson, L., Byrne, T., Johansson, E., Trygg, J., & Vikström, C. (2013). *Multi-and megavariate data analysis basic principles and applications*. Umeå: Umetrics Academy.
Fiske, J. (2010). *Introduction to communication studies*. Oxfordshire: Routledge.
Forbes. (2015a). Narrative science partner. *Forbes Magazine*. Retrieved March 1, 2015. http://www.forbes.com/sites/narrativescience/
Forbes. (2015b). Intel profit expected to slip. *Forbes Magazine*. http://www.forbes.com/sites/narrativescience/2015/07/13/intel-profit-expected-to- slip/
Freytag, G. (1896). *Freytag's technique of the drama: An exposition of dramatic composition and art*. Chicago, IL: Scholarly Press.
Gerrig, R. J. (1993). *Experiencing narrative worlds: On the psychological activities of reading*. New Haven: Yale University Press.
Gershon, N., & Page, W. (2001). What storytelling can do for information visualization. *Communications of the ACM*, *44*(8), 31–37.
Green, M. C. (2004). Storytelling in teaching. *APS Observer*, *17*, 4.
Green, M. C., & Brock, T. C. (2002). In the mind's eye: Transportation-imagery model of narrative persuasion.
Haroz, S., Kosara, R., & Franconeri, S. L. (2015, April). Isotype visualization: Working memory, performance, and engagement with pictographs. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems* (pp. 1191–1200). ACM.
Heer, J., & Bostock, M. (2010, April). Crowdsourcing graphical perception: using mechanical turk to assess visualization design. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 203–212). ACM.
Howson, C. (September 4, 2014). Visual data discovery: 4 Storytelling approaches compared. *Information Week*.
Hullman, J., & Diakopoulos, N. (2011). Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, *17*(12), 2231–2240.
Hullman, J., Drucker, S., Riche, N. H., Lee, B., Fisher, D., & Adar, E. (2013). A deeper understanding of sequence in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics*, *19*(12), 2406–2415.

Intel Corporation. (2015). Intel website. Retrieved 1, 2015, http://www.intel.com/content/www/us/en/homepage.html
Keim, D. A. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, *8*(1), 1–8.
Kock, N. (2004). The psychobiological model: Towards a new theory of computer-mediated communication based on Darwinian evolution. *Organization Science*, *15*(3), 327–348.
Kwon, B. C., Stoffel, F., Jäckle, D., Lee, B., & Keim, D. (2014). Visjockey: Enriching data stories through orchestrated interactive visualization. In *Poster compendium of the computation + journalism symposium* (Vol. 3).
Kydd, C. T., & Ferry, D. L. (1991, January). Computer supported cooperative work tools and media richness: An integration of the literature. In *Proceedings of the twenty-fourth annual hawaii international conference on system sciences* (pp. 324–332). IEEE.
Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, *11*(1), 65–100.
Lee, C. E. (2010). Face-to-face versus computer-mediated communication: Exploring employees' preference of effective employee communication channel. *International Journal for the Advancement of Science & Arts*, *1*(2), 38–48.
Lengler, R. (2006, July). Identifying the competencies of'visual literacy'-A prerequisite for knowledge visualization. In *Tenth International Conference on Information Visualisation (IV'06)* (pp. 232–236). IEEE.
Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., & Byers, A. H. (2011). Big data: The next frontier for innovation, competition, and productivity.
Mayer, R. E., & Anderson, R. B. (1991). Animations need narrations: An experimental test of a dual-coding hypothesis. *Journal of Educational Psychology*, *83*(4), 484.
Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*(2), 312.
Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology*, *86*(3), 389.
McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review*, *90*(10), 60–68.
McCloud, S. (1993). Understanding comics: The invisible art. Northampton.
Munzner, T. (2009). A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, *15*(6), 921–928.
Munzner, T. (2014). *Visualization analysis and design*. Natick: AK Peters/CRC Press.
Narrative Science. (2015). Storytelling is the last mile in big data and analytics. Narrative science website. Retrieved March 1, 2015. http://resources.narrativescience.com/h/i/124944227-what-is-natural-language-generation
New York Times. (2013). 2013: The year in interactive storytelling. *New York Times Newsgraphics.* http://www.nytimes.com/newsgraphics/2013/12/30/year-in-interactive-storytelling
Nichols, B. (2017). *Introduction to documentary.* Bloomington: Indiana University Press.
Oatley, K., Green, M. C., Strange, J. J., & Brock, T. C. (2002). Narrative impact: Social and cognitive foundations. *Narrative impact: Social and cognitive foundations.*
Reiter, E., & Dale, R. (1997). Building applied natural language generation systems. *Natural Language Engineering*, *3*(1), 57–87.
Robertson, G., Fernandez, R., Fisher, D., Lee, B., & Stasko, J. (2008). Effectiveness of animation in trend visualization. *IEEE Transactions on Visualization and Computer Graphics*, *14* (6), 1325–1332.
SAP. (2015). SAP Lumira: Stories. SAP Help Website. Retrieved March 1, 2015. http://help.sap.com/saphelp_lumira120/helpdata/en/1b/df5934709d410ea57d530ab7333fdc/content.htm
Seeking Alpha. (2015). Intel earnings preview: Strong quarter expected. seeking alpha website. http://seekingalpha.com/article/2812135-intel-earnings-preview-strong-quarter- expected
Segel, E., & Heer, J. (2010). Narrative visualization: Telling stories with data. *IEEE Transactions on Visualization and Computer Graphics*, *16*(6), 1139–1148.

Shannon, C. E., Weaver, W., & Burks, A. W. (1951). *The mathematical theory of communication*. Chicago: UI Press.
Shaw, G., Brown, R., & Bromiley, P. (1998). Strategic stories: How 3M is rewriting business planning. *Harvard Business Review*, *76*(3), 41–49.
Shneiderman, B. (2003). The eyes have it: A task by data type taxonomy for information visualizations. In *The craft of information visualization* (pp. 364–371). San Francisco: Morgan Kaufmann.
Shneiderman, B. (2008, June). Extreme visualization: squeezing a billion records into a million pixels. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data* (pp. 3–12). ACM.
Slate. (2014). The first news report on the L.A. earthquake was written by a robot. *Slate Magazine*. http://www.slate.com/blogs/future_tense/2014/03/17/quakebot_los_angeles_times_robot_journalist_writes_article_on_la_earthquake.html
Tableau. (2014). 8.2 Preview: storypoints. *Tableau*. http://www.tableau.com/about/blog/2014/5/82-preview-tell-story-your-data-story-points- 30761
The Atlantic. (2012). Can the computer at narrative science replace paid writers. *The Atlantic Magazine*. http://www.theatlantic.com/entertainment/archive/2012/04/can-the- computers-at-narrative-science-replace-paid-writers/255631/
The guardian. (2015). datablog. Retrieved March 1, 2015. http://www.theguardian.com/data
Tversky, B., Morrison, J. B., & Betrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies*, *57*(4), 247–262.
Ward, M. O. (2008). Multivariate data glyphs: Principles and practice. In *Handbook of data visualization* (pp. 179–198). Berlin: Springer.
Yee, K. P., Fisher, D., Dhamija, R., & Hearst, M. (2001). Animated exploration of dynamic graphs with radial layout. In *IEEE Symposium on Information Visualization, 2001. INFOVIS 2001*. (pp. 43–50). IEEE.
Yi, J. S., Kang, Y., & Stasko, J. (2007). Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization and Computer Graphics*, *13* (6), 1224–1231.
Yseop. (2015). What is smart BI. Yseop website. Retrieved March 1, 2015. http://yseop.com/EN/smart-business-intelligence-why
Zacks. (2015). *Earnings preview*. Zacks Website. Retrieved March 01, 2015, http://www.zacks.com/earnings.
Ziemkiewicz, C., & Kosara, R. (2008). The shaping of information by visual metaphors. *IEEE Transactions on Visualization and Computer Graphics*, *14*(6), 1269–1276.