A Framework for Analysing Casual Data Visualisations as Narrative Media

Neil O'Carroll

A research paper submitted to the University of Dublin, in partial fulfilment of the requirements for the degree of Master of Science Interactive Digital Media

2015

Declaration

I declare that the work described in this research paper is, except where otherwise stated, entirely my own work and has not been submitted as an exercise for a degree at this or any other university.

Signed: ____________________________

Neil O'Carroll

15 May 2015

Permission to lend and/or copy

I agree that Trinity College Library may lend or copy this research paper upon request.

Signed: ____________________________

Neil O'Carroll

15 May 2015

Acknowledgments

I would like to sincerely thank Dr Declan O'Sullivan for his invaluable critique, guidance and encouragement throughout the research and writing process. In addition, I would like to thank Dr Glenn Strong and all the lecturing staff on the Interactive Digital Media course for their enthusiasm and patience in bringing the class through both teaching terms. Finally, my classmates, friends and loved ones deserve due praise for tolerating my serial absence and trying presence. Your support has been critical in this process.

Abstract

In recent data analytics research, much has been written about the function of data visualisations as storytelling media that improve comprehension of large and complex datasets for both expert and non-expert users. Concurrently, the availability of consumer applications that allow for easy manipulation and display of data has given rise to what is known as casual data visualisation. These casual visualisation systems bring new applications of data to new audiences, and so research into the nature of data visualisations in casual modalities must be carried out. This paper investigates the storytelling opportunities in the new paradigm of casual data visualisation by devising and applying a framework for analysing casual data visualisations as narrative media. In this research paper, I first establish the context of casual data visualisations in the era of Big Data before investigating the narrative dimensions of casual visualisation systems. Through this methodological analysis, I devise and define a taxonomy of the narrative dimensions of casual data visualisation. In the final chapter, I analyse three casual data visualisation systems by applying the devised framework. The systems are selected as applications that represent three interesting areas of casual data visualisation: online cultural database (Rap Stats), lifelogging (Reporter) and casual tools for exploring social media data (YouTube Trends). Opportunities and barriers for successful narrative in each system are identified through the analysis, and my conclusions demonstrate the functional utility of my framework for designers, researchers and critics engaging with casual data visualisations as narrative media.

Table of Contents

Introduction

Data visualisations have enjoyed a surge in popularity among casual users that has been propagated by the development of non-specialist systems, giving rise to what has become known as casual data visualisation (Pousman, Stasko & Mateas, 2007). The use of charts and graphs as tools of corporate communication by professionals has been eclipsed by the emergence of easily attainable casual data; and so data has become a factual language for journalists, hobbyists and online media users (Madhavan et al., 2012). In this new casual paradigm, data is used to tell stories, and understanding the narrative properties of data visualisation becomes a new challenge for research (Pfannkuch, Regan & Wild, 2010; Rosling, 2013). This research paper presents a framework for analysing casual data visualisations as narrative media. As systems on the boundary of computer science, art and design; casual data visualisations remain unclaimed by any research field, and so new methodologies must be developed for design, analysis and understanding in this new paradigm (Ehrmann, 1995; Moore, 1998). Devising a framework for engaging with these systems as narratives is an important step for casual data visualisation research.

This paper outlines the process of devising and applying a framework of narrative dimensions for analysing casual data visualisations. The first chapter examines the context of this research, arguing that Big Data technologies as well as new cultural attitudes to data have given rise to casual interactions with data in everyday computing and everyday life. The second chapter explores narrative and explicates methodological definitions of the narrative dimensions of casual data visualisations. Finally, the third chapter applies these analytic dimensions to three casual data visualisations systems in order to uncover the narrative salience of each system and test the validity of the framework itself. The results of this sample application demonstrate the need for the development of new analytic approaches to the emerging field of casual data visualisations.

1 Big Data and the Emergence of Casual Data Visualisation

1.1 Data for Everyone

This chapter will establish the position of casual data visualisation in the current climate of digital media and computer technology. This paper proposes that casual data visualisation is just one exciting application of data as it has evolved into a raw material for creative arts, science, commerce and research. In order for this paper to successfully explore data visualisation, it is necessary to first assess the modern digital landscape; in particular the collection, storage and processing of quantified information. The following section will identify the technological and social changes that have allowed for the emergence of new data paradigms in order to establish the theoretical context within which data visualisation will be examined in detail.

1.1.1 What is Big Data?

Data is the term used to refer to the contents of organised stores of information. The current pace at which data is gathered, quantified and measured by computer systems has led to the emergence of a new field of research and industry: Big Data. While to some extent a buzzword, Big Data is certainly a useful term to discern interactions with large, complex datasets that would be difficult or impossible to process using traditional applications. The distinction of this new paradigm from traditional statistics and data analytics is characterised by many factors, and the social implications of Big Data are derived from new insights achieved through the field. Eaton et al. (2012) distinguish data above a threshold of volume, variety and velocity as 'big', while Zaslavsky, Perera & Georgakopoulos (2012) attribute the increase of ambient data; the quantifiable information left behind by users interacting with the Internet of Things[1], as a major contributing factor to the phenomenon. In any case, many academics, corporate researchers and business leaders have attested to the rapidly growing interest in the efficiency at which everyday information is handled by computer networks and, crucially, the development of rich economies around the troves of personal data accumulated online (The Economist, 2010).

1.1.2 Technology, Society and Data for the People

The emergence of this era of Big Data, wherein the relationship between society and digital information is fundamentally changed, has not been brought about by the advancement of computer technologies alone. It is, rather, factors including the widespread adoption of new data-enabling technologies and practices, changing attitudes of data engagement and the advent of innovative data-driven artefacts that allow digital industries to approach new frontiers of understanding through data

[1] the emerging move towards a standardised, unified protocol for connecting everyday objects in order to make an intelligent grid of connected devices that is seamlessly integrated with the physical world (Barbry, 2012; Saint, 2015)

(Kosciejew, 2013). This section will trace the contribution of technological development and changing social attitudes to new applications of data for users of all types.

As of 2015, with recent assessments of top analytics company revenues in the hundreds of millions (Information Management, 2015), and emerging analytics markets in countries such as India estimated to be worth billions (Amberber, 2015); it is evident that the Big Data industry is in a state of prosperity. However, the exponential development of computer technologies and their applications can obscure perspective on what is a mature discipline and what is in its infancy. Just as the mobile phone evolved from the affordable communications device of the 1990s to a ubiquitous personal computer interface by the early 21st century (Giachetti & Marchi, 2010), so Big Data is perhaps only beginning to reveal its potential for economic and social advancement. What implications does this have for the average citizen looking to gain from all this innovation?

By some measurements, the current data ecosystem is fragmented and inefficient, and in many cases the risks and liabilities for companies and individual users engaging with large-scale data can exceed the potential returns (World Economic Forum, 2011). This begs the question from the perspective of users: what can interactions with data offer that are worth the risks posed to personal privacy and autonomy? Does high-velocity data interpretation offer any great benefits for the individual, or is it just a covert means to a marketing end? Symptoms of technophobia begin to appear as one moves closer to the root of society's data fixation. Indeed, it has been claimed that the actual contribution of technology to the Big Data phenomenon is only superficial, as the field is driven at its core by humanity's innate desire to understand the universe it inhabits (Cukier & Mayer-Schoenberger, 2013). This perspective is enlightening, but is arguably true of all technology and the duality of creator and creation: the digital chicken and egg. Technology and society are inextricably linked, and rationalising how the exabytes of humanity's everyday data are understood is just another question raised in the long history of this coupling. Measurable contributing factors to the new data paradigm identified by Cukier & Mayer-Schoenberger in the same paper are more focussed on *how* the data is interpreted by humans rather than the existence of large databases and complex algorithms themselves (ibid.). The new attitudes identified by Cukier & Mayer-Schoenberger that distinguish Big Data from historical statistical practices are:

1. The lack of necessity to project hypothetical figures from small samples of data
2. The preference for larger volumes of data over highly curated data
3. The lack of need to understand the reasons for correlation in data, but rather the preference to recognise that there is a correlation

Interactions with data under these assumptions allow for new insights by analytics experts and professionals, but this paper more specifically concerns the interactions of non-experts with current data systems. One key example of such interaction is lifelogging; the practice of recording and personally documenting daily and mundane activities.

1.1.3 Lifelogging and Human-Data Interactions

Lifelogging is a digital and social media subculture that has evolved around the abundance of data-enabling technologies[2]. Quantifying and recording everyday data such as diet, location and mood can empower users towards self-expression and self-actualisation. Journalists and artists have become early adopters of lifelogging, with prominent work in the field including the obsessively diligent *Feltron Annual Report* (Felton, 2015a), a self-published dossier of personal data by visual designer Nicholas Felton; and intimate artwork such as *Images of the Artifacts Used by the Main Hand* (Frigo, 2015), an ongoing photography project by Alberto Frigo, who will continue documenting everything he holds in his right hand for the next 25 years.

Lifelogging is notable as a manifestation of embodied interaction; the theoretical approach to human computer interaction (HCI) that recognises computer systems as systems embedded in social meaning, especially meaning generated through the analysis of the mundane (Dourish, 2001). Though not yet a mainstream practice, lifelogging directly engages non-professional and non-technical users with Big Data and provides a useful case study for rationalising human-data interactions (HDI): a recently emerging research area that explores social and psychological perspectives on Big Data. In their paper on HDI, Mortier et al. (2014) pose many questions about the nature of the human-data relationship discussed in this chapter. They aim to develop criteria by which to qualify and quantify HDI, and they begin by addressing the issue of personal data. Personal data can refer to data both about an individual or created by an individual (ibid.). The quantification of personal data in lifelogging provides insight into the nature of casual data interactions that lead to expressive or creative applications of data. However, with the recognition of this new discipline comes the recognition of new challenges. Without employing large and expensive training schemes to realise new levels of data understanding, how can society be motivated to engage with the historically dull and daunting subject of numbers and statistics? In addition, the distrust and aforementioned technophobia of the data age is here acknowledged, but a move towards transparency of mythical algorithms is at the loss of valuable intellectual property for the companies spearheading innovation in the data era (ibid.). Most critically, HDI research desires to place humans at the centre of the data economies they already occupy; a

[2] Adapted portions of the brief lifelogging research presented here have been submitted as supplementary, secondary research documentation for a games design document for the 2015 Interactive Digital Media course at Trinity College Dublin. A statement of the primary use of the material for this research paper was also submitted at the time of the assignment.

democratic stance against the possibility that Big Data becomes void of influence from society and remains a shrouded, misunderstood mathematical voodoo disproportionately weighted in favour of data algorithms, aggregators and analysts.

1.1.4 Addressing Casual Data Interactions

The threat of the autonomous Big Data machine becoming an oppressor of society is counterbalanced by the view that the very same data will serve to empower users who engage with it. Young (2012) proposes that individuals who use digital information to create large virtual datasets about themselves can positively alter their realities in the physical world; and that data offers agency to users who interact with it rather than posing a threat to personal privacy or sovereignty. This paper proposes that casual users of data are those users who seek these empowering interactions, and that the scope of these interactions can range from quantifiable mundane activities such as diet, location or communication (personal data); to external but potentially empowering datasets including sports statistics, historical language patterns or data used to support news reports (cultural data). Casual data is understood as this personal and cultural data. However, access to casual data by casual users of data may be inhibited by hangovers from older practices of statistics as a mathematical science. Prevailing challenges of government policy, bureaucracy and deliberate use of complicated mathematical processes can hinder the realisation of public access to unambiguous, empowering data (Rosling, 2013). The wider availability of this kind of data serves to support movements such as lifelogging in resolving the dissonance between Big Data technology and users of casual data. Promoting casual data interactions is a step towards persuading citizens to accept a fact-based view of the world (ibid.), and there is certainly a critical interdependence between social attitudes, data policies, technology and, ultimately, salient data interactions.

1.2 Visualising Information

Data visualisation is one such application of data interaction and is a medium that approaches solutions to many of the challenges proposed by HDI and Big Data research discussed in the previous section. Data visualisation will therefore be defined and analysed in this section with the aim of establishing its relevance to human-data interactions. Casual data visualisation will subsequently be identified as a sub-genre of data visualisation and rationalised as an agent of sensemaking, information transparency, analytics and; most critically, a catalyst of ‘data-to-knowledge’: the potential for society to exploit the wealth of information offered by Big Data (Mortier et al. 2014).

1.2.1 What is Data Visualisation?

Visualisation concerns the mapping of discrete numerical data to visual representations: the translation of information modalities into image modalities (Manovich, 2011). The applications of visualisation, like data, range from scientific use by data experts to artistic use by design experts, though preferred visual techniques differ between the two disciplines (ibid.). This chapter will later explore the use of visualisation by casual users who are experts in neither field, but wish to exploit visualisation for its empowering ability to reduce cognitive load and communicate concisely when engaging with casual data.

Displaying data in a visually intelligible way relies on two core principles that are manifest in all celebrated examples of data visualisation since its beginnings in the work of William Playfair (Tufte, 1986). The first principle is reduction: the scaling of large and unmanageable information series to human-readable visual forms. The second principle is the use of spatial variables: area, length, shape and position to not only display data meaningfully but also access realms of artistic and individual expression. The most basic function of visualisation is to achieve understanding by faster and simpler means than looking at the vast arrays of numbers that visualisations are devised from (CACM Staff, 2014). Fundamentally, data visualisation offers users the opportunity to overcome barriers that are inherent to the sciences by bypassing the semantic and potentially unclear mathematical meaning of numerical or linguistic data and introducing instead a method of information communication that is uniquely defined on a per-instance basis. Each data visualisation is, ideally, considered a single-instance of communication and should not rely on previous knowledge gained from interactions with other visualisations. For example, chemistry textbooks will feature figures denoting a number of electrons orbiting the nucleus of an atom. Basic visualisations coupled with a simple key to explain which subatomic particle is which enable the near-immediate understanding of hundreds of years of experimentation and hard scientific research. The following sections concern the extension of this principle of information representation to augment the interpretation of casual data.

1.2.2 Casual Data Visualisation

Recalling the explication of casual data as personal or cultural data that has the potential to empower individuals towards agency in the modern Big Data era; casual data visualisation, then, is any visualisation that represents this kind of data and can be accessed by users through easily available tools and platforms. Casual data visualisation strives to amplify cognition through the artistic representation of everyday data; data and its technologies are objects of reflection, and visualisations encourage repeated interaction by a large population of users with the aim of developing or expanding on a personal relationship with the dataset (Pousman, Stasko & Mateas, 2007). This paper's supposition that casual data forms the basis of casual data visualisation incorporates the notion that there is a vested interest in engaging with the data, so long as this interest is not related to work or professional engagement with data visualisation. Casual interactions with data visualisation are a subset of data visualisation but, much like lifelogging, include systems and media on the boundaries of data science, ubiquitous computing, design, art, popular psychology and computer programming (ibid.). Three systems of casual data visualisation will be analysed in the third chapter of this research, selected as representative applications of the emerging field discussed in this chapter.

1.2.3 Optimising Casual Use Cases

In order for casual data visualisation to further enhance human-data interactions in the future, a better understanding of the motivating factors for casual users to engage with data must be realised. Sprague & Tory (2015) explore how people engage with visual representations of data in casual contexts and consolidate existing research to identify five key factors that influence motivation to interact with casual data visualisations:

1. **Usefulness:** does the visualisation serve to help the user?
2. **Self-Reflection:** can the user learn something about themselves through the visualisation?
3. **Learning Costs:** is there a steep learning curve?
4. **Personal Interest:** does the user have an existing relationship with the data?
5. **Social Interaction:** can the visualisation generate shared experiences, community involvement or increase social interaction?

Where these factors have been considered, prolonged and repeated interaction with casual data visualisations was observed; the conclusive findings of the research posit the Promoter-Inhibitor Motivation Model (PIMM), where continued interaction occurs when perceived costs are outweighed by perceived benefits when a user is considering engaging with casual data visualisation (ibid.). Being aware of these motivating factors in designing casual data visualisations could spur the trend towards adoption of emerging data-enabling technologies, increasing engagement with casual data and empowering users through interaction with data. These motivating factors inform the analytic dimensions of casual data visualisation devised in the second chapter.

1.3 Chapter Conclusions

The significant conclusion of this chapter’s analysis is the dichotomy of Big Data as an incomprehensible entity and casual data visualisation as an agent of enlightenment. Those who feel daunted by the thought of traversing large tables of data to discover information about weather, bus routes or television schedules are perhaps unaware of the work that casual data visualisation has done to make these tasks easier through the advent of now-common interfaces and data displays, but there remain many other data interactions that are ripe for improvement through visualisation. This chapter has established the role of casual data visualisation in the dynamic environment of human-computer and human-data interaction. Data visualisation is a product of the Big Data phenomenon whose origins and influence have been explored as part of this chapter’s rationale. Casual data visualisation is an emerging sub-genre of data visualisation that incorporates many diverse and novel applications across a range of disciplines. Beyond research applications, casual data visualisation may be a critical proponent of new social attitudes towards emerging technology paradigms ranging from ubiquitous computing to concepts of true data-driven lifestyles.

In contrast to the analysis in Chapter 1, the next chapter concerns narrative; the other core concept of this dissertation. The second chapter will first look at perspectives on narrative before incorporating prior analysis of casual data visualisation to consolidate the research into a single methodology. The outcome of the second chapter will provide a framework for the casual data visualisation analysis of the third chapter.

2 Exploring Data Visualisations as Narrative

2.1 Data Narratives

Realising the potential of data visualisation for casual users necessitates the understanding of data and database as narrative artefacts: means of telling stories and communicating information about the human experience. The second chapter of this research paper explores narrative. It is impossible to navigate this analysis without steering the tone towards philosophical and sociological methods in order to fully appreciate the position of database, narrative and data visualisation as objects that populate broad human experience. Moore (1998) advocates this repositioning of statistics among the liberal arts in order to promote innovation in the field and ensure the technology and theory of mathematics is not accelerated past relevance to humanity and society. In light of this, the rationale in my mode of analysis considers humans at once as *users* by virtue of their interaction with databases; and *characters* in the narratives of life.

This chapter will first define narrative and investigate its relationship with the database. Perspectives on narrative theory will be identified and, by then examining data visualisation through the lens of narrative theory; useful analytic language and methodological parameters will be generated. The terminology and theoretical perspectives identified in this chapter will form the basis of the main research analysis in the third chapter, and therefore the goal of this chapter is to develop coherent analytical instruments via the exploration of data-driven narratives. Where identified, useful terminology will be displayed in italicised font for later use and discussion.

2.1.1 What is Narrative?

> *"The narratives of the world are numberless. Narrative is [...] a prodigious variety of genres, themselves distributed amongst different substances - as though any material were fit to receive man's stories"*
> (Barthes and Heath, 1987, p.87)

In his *Introduction to the Structural Analysis of Narratives*, Barthes relates narrative to substance, story and man; three keystones of communication (ibid.). The basic act of communicating information through any media fundamentally requires the presence of the information itself; the physical communicator and, crucially; the method of communication. The communicator and information criteria in this system have been discussed under the guises of *human* and *data* in the first chapter of this research paper; it is the concern of this chapter that the *method* in which information is communicated is thoroughly understood. Narrative theory enables this understanding, as it does not consider the mere existence of event, information, substance or data as analogous to story; but rather positions metatextual elements including the structure of the telling of the story[3], the audience, the storytelling medium[4] and the storyteller as central to the eventual meaning of the text.

At its most basic, narrative can be defined as a representation of events (Abbott, 2002, p.12; Chatman, 1978, p.19). The events are unambiguous entities inherent in everyday life; their representation through the technologies of our history, however, is the cause of contention in the understanding of narrative. With each new era of media and technology, the ontology of narrative becomes changed and complicated. Resolving this complication is a primary goal of the wider field of communication studies, and so this paper specifically prioritises the understanding of narrative as it relates to Big Data technologies and casual data interactions. Narrative is neither magnificent literature nor extravagant art; it is the necessary protocol of transmission for the everyday events of life and, as the everyday becomes further quantified in the era of Big Data, so must a better understanding of the narrative opportunities of the database be realised. Taking influence from Barthes succinct, if not somewhat opaque, introduction to narrative quoted in previously; this paper defines narrative as the resulting human experience of story found at the meeting place of data and medium. As narratives are used to tell stories; so 'storytelling' is here used to describe the act of presenting information in a cognitively efficient way.

2.1.2 Database and Narrative

From oral traditions of storytelling to literature, film, art and music; the artefacts and technologies of human history transform how information is communicated, preserved and understood. The scale, speed and scope of Big Data has been identified in this paper as an enabling factor in affecting greater human interaction with information as quantified data, but what effect does the access of society to colossal databases of personal and cultural data have on narrative? Investigating the dialectic of database and narrative recognises that the human experience is at the epicentre of both concepts. The database serves to quantify, analyse and automate aspects of society in order to benefit industry, commerce and communication; the narrative serves to recount the events and experiences of the participants in this society. The database, powerful in its capacity to store and identify relational events, needs narrative to give its contents meaning (Freedman et al., 2007; Pfannkuch, Regan & Wild, 2010). Database and narrative are, by this measure, both halves of the ontology of the world;

[3] commonly understood as the communication of any information from one person to another; *"an account of real or imagined events"* (Dictionary.com, 2015); *"a statement regarding the facts pertinent to a situation in question"* (Merriam-webster.com, 2015)
[4] in broader communication theory which falls outside the scope of this brief discussion of narrative theory, it has been popularly proposed and theorised that the method of communication is its own information and the two concepts are so intertwined they are inseparable: *"the medium is the message"* (Mander, 1978; McLuhan and Lapham, 1994)

two interdependent solutions to the same problem of we humans, as a species, documenting and understanding our existence (Manovich, 2007).

This view of the entirety of life as a binary of narrative and database, however, is restrictive in its simplification and parameterisation of the intricate nature of how humans exist in the world. It is perhaps more useful to view narrative and database simply as two instruments of understanding. No one story can provide a complete and accurate account of a real and complex event (Weinbren, 1997); and, conversely, there is no single optimal way to communicate the information in a given database. Certainly, database reconfigures narrative and the possibilities of human storytelling (ibid.), but the concept of addressing factual data as subjective sources of information is not new to the era of Big Data or even modern computing. Rather, unlocking the meaningful, interesting narratives in cold, mathematical data has been central to good data analytic practice for at least four decades; despite not strictly using the now-familiar terminology of literary and communications theory (see Ehrenberg, 1975; Tukey, 1977).

The litany of multidisciplinary research into database and narrative requires focus in order to generate useful analytical methodology for this paper. It is necessary at this stage to adapt the most relevant work in this area in order to define our own parameters of applied theory. For this research, the database is identified as a narrative artefact in recognition of:

a) the reshaping of data into a form of storytelling (Bass, 1999; Klein, 2007; Weinbren, 1997)
b) the propagation of data storytelling as a product of the identified trend towards casual data interactions in everyday human experience (see Chapter 1)

Data narratives are just another invented possibility in the ongoing development of new methods of documenting and recounting mundane or fantastic events, and will enjoy their position as exciting and inspiring narrative objects; the underlying and timeless function of technology as a means of communicating stories allows for successful investigation of data narratives from perspectives adapted from general practices of storytelling. Nietzsche saw story as the sculpting of madness (Klein, 2007). The next section of this chapter asks of the madness of data narratives in contemporary digital culture in further pursuit of the methodological parameters of this research: how do we tell successful stories through data visualisation?

2.2 Dimensions of Narrative Data Visualisation

Data visualisations have been identified in the first chapter as media objects on the boundaries of mathematics, art and technology. Representing quantified data in a visually coherent way can be as simple as colour-coding a display of numbers or as complicated as mapping the geographical movement of populations; it is the stories that are told through these visual representations of data that are the focus of this section. Document analysis of research in the area of storytelling through data visualisation has uncovered a broad range of perspectives on the functions and components of narrative data visualisation, and the simultaneous emergence of casual data interactions has given researchers a new lens with which to examine data narratives. These two research areas have informed the categorical division of the following analysis. In order to investigate opportunities for successful storytelling through the medium; narrative data visualisation will be discussed under the headings of *Genre*, *Structure*, *Platform*, *Stylistic Embellishment*, *Interactivity* and *Social Collaboration* (see Table 1). These dimensions of narrative data visualisations have been identified by this author as defining components of the medium, and aim to provide a framework for analysing casual data visualisations as narrative media.

Table 1 on the next page shows the concepts incorporated in each dimension of narrative visualisation analysis. The method of devising these six dimensions involved reviewing, analysing and coding literature pertaining to data visualisation as narrative and data visualisation for casual use cases. Correlating similar functional language between discrete research in both fields, and cross-referencing the semantic and technical interpretations of common terms gave rise to the six dimensions that will be examined in detail in this chapter.

As well as listing the key concepts and notable research influences for each dimension of analysis, the table notes where a term in the framework has been adopted directly from research; where a common term has been defined in the context of the framework; and where a related group of existing concepts has influenced analytic terminology devised specifically for analysing an identified narrative dimension of casual data visualisation.

2.2.1 Genre

Genres are useful subdivisions of media that group individual works together based on shared characteristics. Knowing the intended genre is helpful for informing the decisions of both audience and creator before engaging with a particular work. Film and literary genres are distinguished by form and aesthetic rather than content; a story concerning a relationship affected by death could be presented as romance, thriller, horror or comedy. In data visualisation, this separation of content and presentation offers exciting opportunity for creators as it suggests that even the most dense and mundane data can be made into salient and successful narrative through informed design. Loosely, visualisations could perhaps be divided into genres based on technology, levels of interaction or subject matter, but this is unhelpfully vague and counterintuitive to genre models in other narrative media.

Dimension	Incorporated Concepts	Research Influences	Source of Terminology
Genre	Number of Discrete Charts (Frames) Order of Presented Frames Type of Text Supporting Chart	Figueiras (2013) Segel & Heer (2010) Sprague & Tory (2012)	Adopted from Segel & Heer (2010) and Figueiras (2013)
Structure	Control of Narrative Flow Reader-Driven or Author-Driven Narrative Logical Order of Narrative Transitions Between States	Hullman et al. (2013) Segel & Heer (2010)	Common term defined here in the context of narrative data visualisation
Platform	Hardware & Software Technology User Interface Interaction Pattern Potential Interactivity	Childs et al. (2013) Roberts et al. (2014)	Original terminology derived from cited research
Stylistic Embellishment	Visual Design Aspects Imagery Colour Non-data Ink	Bateman et al. (2010) Vande Moere et al. (2012) Viégas & Wattenberg (2007)	Original terminology derived from cited research
Interactivity	User Agency Interface Manipulation Engagement Intended Narrative Outcomes	Elmqvist et al. (2011) Madhavan et al. (2012) Satyanarayan, Wongsuphasawat and Heer (2014)	Common term defined here in the context of narrative data visualisation
Social Collaboration	View Sharing Doubly Linked Discussion Asynchronous Discussion	Heer, Viégas & Wattenberg (2007) Mackinlay (2009) Sprague and Tory (2015)	Adopted from Heer, Viégas & Wattenberg (2007), Mackinlay (2009) and Sprague and Tory (2015)

Table 1 - Dimensions of Narrative Data Visualisation Analysis

Formalising tropes of genre more specifically is necessary to further understanding in this emerging medium. Segel & Heer (2010) document a comprehensive analysis of narrative data visualisations using examples from news media, academic research and hobbyist communities. From this research, seven genres of narrative visualisation are identified: *Magazine Style, Annotated Chart, Partitioned Poster, Flow Chart, Comic Strip, Slide Show,* and *Film/Video/Animation* (ibid.). These genres are distinguished primarily by the number of frames (discrete charts of data) and the order in which the frames are presented, but also differ in their use of visual tropes such as descriptive text or annotation. Visualisation genres are not mutually exclusive and, much like the genre divisions of cinema, music and literature, are often combined in visualisations to create interesting narratives. The ability to place a visualisation in one of these genres is a first step in developing useful language for public discourse, and Segel & Heer's paper has been widely cited as fundamental to the formalisation and theorisation of the emerging field (see Hullman et al., 2013; Kosara & Mackinlay, 2013; Sprague & Tory, 2012). Elsewhere, further work to devise a typology of data visualisation has identified genres that are unique to online visualisations such as *Tag Cloud* and *Game*; as well as more generic visualisation forms such as *Map* and *Chart/Diagram* (Figueiras, 2013). As such, it is not the aim of my research to complicate or redefine these genres; the above narrative visualisation types will be used to preliminarily distinguish and categorise the casual data visualisation applications analysed in the third chapter.

2.2.2 Structure

Popular narratives in storytelling media such as films and novels tend towards linear structures. These stories aim to be logical, complete and unambiguous (Weinbren, 1997), but linearity does not represent or integrate with the stop-start interaction patterns of humans experiencing the world. With a broader range of possible structures, data visualisations offer storytelling opportunities that enable new understanding beyond mathematics and statistical science (Ehrmann, 1995; Haddadi et al., 2013; Manovich, 2011). Before even considering interactive possibilities; static visualisations on paper present many options for user exploration: single-frame charts, multi-frame charts in an order or multi-frame charts with no defined order. Access to discrete, detailed content and the ability to compare, remove or save states of the visualisation are afforded by dynamic and interactive story structures (Heer, Viégas and Wattenberg, 2009; Madhavan et al., 2012).

Structure is linked to genre but not necessarily dictated by it, and the structure of a visualisation influences how much of the narrative is driven by the author and how much is in the control of the user (Segel & Heer, 2010). Optimising the position of a narrative data visualisation on the spectrum of reader-driven to author-driven story is crucial in order to successfully communicate the data holistically and memorably for the user (Pfannkuch, Regan & Wild, 2010). Hullman et al. (2013)

document an empirical survey of the elements of structure in a large corpus of data visualisations with an aim to provide designers with informed choices in selecting the sequential or freeform structure through which a visualisation is presented. *Causal*, *Temporal* and *Comparative* transitions between states of data visualisation display along with the ability to control the *Level of Detail* and the *Spatial Proximity* of visual elements are identified as structural factors that influence the salience of the final narrative (ibid.). Elsewhere, these spatiotemporal and interactive aspects of structure have been reduced to three general structures found frequently in online applications: *Interactive Slideshow, Drill-Down Story* and *Martini Glass Structure* (Segel & Heer, 2010). The Martini Glass Structure, where the user is first guided through aspects of the story before being allowed freer interaction with the visualisation, is identified as the most common among surveyed examples and noted as an archetypal structure of the emerging narrative medium that balances author and reader-driven story elements in a clear and intuitive way (ibid.).

Activating logical and creative responses to large datasets that are otherwise opaque is perhaps the most fundamental function of narrative data visualisation, but the factual data at the heart of the story must not be obscured in designing the narrative (Gershon & Page, 2001; Rosling, 2013). For this reason, balancing familiar structural story elements with new, unexpected narrative components unique to the medium is an example guideline formula for successful narrative data visualisation design. For instance, a visualisation should first define the context of the data: the location and time of event information; this is analogous to the so-called 'establishing shot' in filmmaking, where the story is given a setting from which the audience can logically infer continuity in the details of the events introduced subsequently. The affordances of interactive digital media applications enhance the established logical narrative structure by allowing the user to emphasise, review or recontextualise elements of the story at their discretion (Wohlfart & Hauser, 2007). By adhering to and augmenting structural guidelines in this way, narrative data visualisation can activate new understandings of complex data without alienating or disengaging the user. Interactivity as a dimension of narrative data visualisation will be discussed in more detail under its own heading later in the chapter.

2.2.3 Platform

The platform on which a visualisation is displayed influences the narrative and can inform other dimensions of data storytelling including the genre, the structure and interactivity. As noted above, static visualisations on paper are not tethered to a particular genre or structure but they are certainly more limited than interactive digital media visualisations in communicating innovative, engaging and memorable narratives. Integrated text and visuals offer more opportunity for clarity and detail than either storytelling medium in isolation (Gershon & Page, 2001); and the scope of narrative visualisation encompasses platforms as diverse as projection mapping, 3D printing, physical computing and multisensory applications (Roberts et al., 2014). For this research paper, the visualisation platform concerns the digital media hardware and software technology through which a narrative data visualisation is presented to the user.

As narrative data visualisations are a consumer medium, the user's needs must be prioritised when selecting a platform for visualisation display. Using familiar genres and traditional storytelling structures on new media platforms has the advantages of creating user-intuitive narratives, but inherent reader biases such as left-to-right reading styles can negatively affect the exploration of more innovative digital media stories (Segel & Heer, 2010). For this reason, designers of narrative data visualisations must consider whether the time spent introducing a user to a new structure or genre is worth the time taken away from their exploration of the narrative, and these decisions are influenced by the visualisation platform. Some established fundamental user requirements of technology platforms in the realm of digital storytelling are low-latency interaction and response time; consistent and representative samples of data at micro and macro levels of display; and actively updatable data, if it is relevant to the visualisation (Madhavan et al., 2012). These affordances of the chosen visualisation platform should aim to be consistent across touch and traditional GUI[5] devices, while ever-evolving consumer technologies offer further opportunities for salient data stories as well as introducing challenges.

Most popular journalistic narrative data visualisations are intended to accompany traditional articles on news websites (Figueiras, 2013; Segel & Heer, 2010), but the shift from desktop to mobile browsing complicates the design of data stories for diverse audiences. Familiar and useful actions for interacting with data visualisations such as hover are not supported by touch devices. Furthermore, the possible display formats of online visualisations range from HD TV screens in living rooms to low quality mobile screens viewed outdoors where direct sunlight may limit already poor visibility and colour distinction (Roberts et al., 2014). Conversely, mobile-viewable and touchable data visualisations allow users to engage with data stories in an integrated and intuitive way, which supports the move towards embodied interaction models of human-centric computing (Cafaro, 2012; Elmqvist, 2011; Haddadi et al., 2013). The future of narrative visualisation, as the medium becomes more consumer-led and adopts a casual modality, is set to involve further integration with emerging virtual reality standards and the use of appropriated surfaces rather that bespoke screens as platforms of display (Childs et al., 2013; Roberts et al., 2014). In addition, Childs et al. (2013) identify a correlation between the increasingly complex methods of handling huge volumes structured and unstructured data and metadata, and the growing need for visualisation software designers to provide frameworks that shield users from data-processing complexity. It is recognised that this trend is indicative of the changing face of data visualisation platforms, modular solutions at hardware and

[5] Graphical User Interface devices such as the traditional point-and-click, multi-window displays of desktop and laptop computers.

software stages of visualisation architecture design are key to overcoming emerging challenges in the field (ibid.). Above all, visualisation platforms must balance simplicity and integrity in an informed and appropriate way, contingent on the goals of the resulting narrative.

In any case, the technology powering narrative visualisations will be driven by user requirements, and many researchers believe that the next step for storytelling is the standardisation of collaborative methods in visualisation software involving sharable visualisation views, fluid interaction models and asynchronous social commentary on visualisations (Heer, Viégas & Wattenberg, 2009; Roberts et al., 2014). Collaboration in visualisation will be explored in more detail later in this chapter.

2.2.4 Stylistic Embellishment

As an emerging narrative medium that draws as much influence from art and design as it does from statistics and data analytics, modern narrative data visualisation problematises many tenets of traditional statistical graphing and data visualisation. Quintessential guidelines for the effective display of quantitative information emphasise the reduction of chart components that are not essential to the representation of data, and criticise embellishments and non-essential visual content for distracting from the core data (Cleveland, 1994; Tufte, 1986; Tufte, 2006). This minimalist approach to data visualisation aims to maximise the proportion of 'data-ink', labelling any non-data ink in visualisation as 'chart junk' (ibid.). However, the aesthetic scope of narrative data visualisation allows for creative employment of so-called junk to add interest to the narrative and, if correctly implemented, aid memory and retention of information rather than inhibiting the effectiveness of the visualisation as a representation of data (Bateman et al., 2010).

The inclusion of aesthetic components beyond bare minimum data points supports the adoption of narrative data visualisations by casual users as tools of personal insight as well as objects of artistic merit (Pousman, Stasko & Mateas, 2007). Crucially, it is understood that the minimalist approach advocated by data visualisation theorists has generally gone unheeded in popular media applications of visualisations, and the designers' decisions to decorate their work could be influenced by users' preference for embellished charts (Bateman et al., 2010; Zacks et al., 2002). The fact that there is a schism between theoretical recommendation and practical application of data visualisation emphasises the need for further research in this area as it evolves to accommodate casual and non-technical use cases.

Even in progressive approaches to narrative data visualisation design, there is no agreed stance on whether or not embellishments are beneficial to the data story. As with the structural dimension of narrative visualisations, there are levels of appropriateness for embellishment contingent on context. For the most part, it is agreed among researchers that clarity of data should remain unaffected by stylistic choices in visualisation design (Cleveland, 1994; Womack, 2014). This establishes a basis from which more ambitious designers can begin to creatively style their data narratives, and the intended purpose or genre of the narrative may inform their stylistic decisions beyond simply representing the data visually. To this end, Vande Moere et al. (2012) comprehensively evaluate the impact of style on visualisation and find that insight types and levels of interaction are influenced by the stylistic features of the data visualisation; adherence to design norms within visualisation genres affects the communication of fact and meaning and can enable access to deeper understanding of data by the user through different characteristics of insight (Saraiya, North & Duca, 2005; Vande Moere et al. 2012).

Aside from distraction; bias and persuasion are thought to be side-effects of superfluous imagery and embellishment in data visualisation, as the inherent subjectivity associated with these additional design features contradicts the objective nature of data (see Cleveland, 1994; Tufte 1986). However, this assertion is based on the premise that minimalist charts are free from bias in the first place, which has not been proven by research (Bateman et al., 2010; Rock, 1992). Even if the collected data is honest and representative in the first place, the very act of extracting raw data from the database and communicating it through any medium is inherently biased; curation and interpretation of data is a native function of this action (Manovich, 2007; Weinbren, 2007). For consumers of data visualisation narratives, the potential for data stories to be biased and persuasive must be recognised as inherent; it does not devalue the medium as a tool of insight, understanding and salient storytelling.

Animation in data visualisation offers opportunities for captivating audiences and developing complex narratives gradually and coherently, but there is an undeniable element of distraction away from core data trends when animation is not thoughtfully employed (Kosara & Mackinlay, 2013; Robertson et al., 2008). The balance of attracting interest in data narratives with animation and communicating data with clarity is contentious. Engagement with visualisations at a casual level certainly demystifies complex data for a wider audience, but there is a point at which the goal for mass public engagement with data gives way to data narratives being consumed as pure entertainment rather than insight. However, this transformation of purpose and reappropriation of data visualisation as visual art is not necessarily detrimental to the wider field of narrative data visualisation (Viégas & Wattenberg, 2007). Ultimately, designers of narrative data visualisations must consider embellishment as a means through which to reinforce in all users a readiness to accept new data visualisation modalities as this new era of casual data interactions evolves. Identifying the ‘sweet spot’ of factual data framed by captivating imagery and animation is critical for designers of narrative data visualisations in the current paradigm (Inbar, Tractinsky & Meyer, 2007).

2.2.5 Interactivity

Interactivity has been cited previously in this chapter as a component of narrative structure that is found in data visualisations in digital media applications. Interactivity is a catalyst of understanding, where a feedback loop of ideas is generated between user and technology (Eisenstein, 1949; Kiousis, 2002; Manovich, 2001). In data visualisation, interactivity gives agency to the user and enables them to engage with data narratives intuitively, resulting in richer and more personal insights into otherwise complex datasets. Temporal and spatial transitions controlled by the user amplify potential cognition of data; zooming in on details or revisiting previous frames help resolve ambiguities, but too much freedom to explore the data with no inherent motivation may result in lower interaction and engagement (Gershon & Page, 2001; Sprague & Tory, 2015). Higher levels of interactivity correlate with reader-driven narrative structures in visualisation; stories intended for activating discovery and deep insight favour more interaction points[6] in the narrative (Segel & Heer, 2010; Vande Moere et al. 2012). To this end, the designer must incorporate interaction into a data visualisation only if is appropriate to the intended narrative outcomes:

- Is there a particular message the author wishes to communicate?
- Is the user exploring the data for analytic or confirmatory purposes?
- Is the visualisation supposed to incite debate or collaborative understanding?

Approaching interactivity in data visualisation in this way preemptively dissuades designers from overloading the user with choice and focusses the structure of the narrative through careful and considered distribution of interaction points (Kosara & Mackinlay, 2013; Madhavan et al., 2012). However, designing for interaction causes problems for practitioners, even when these narrative goals are considered. Interactivity, though understood by users on a semantic level, is noted as a technologically intangible concept that is difficult to quantify and evaluate in visualisation design (Elmqvist et al., 2011). Simply allowing the user to manipulate the interface does not necessarily enhance the visualisation, though somewhere in the spectrum of interaction models there exists the opportunity for the designer to engage the user in a dialogue with the data and, ultimately, achieve salient narrative (ibid.).

External factors also present barriers to useful and meaningful interactivity in data visualisation. A successful model of interactivity is contingent on technology and its limits; hardware, software and network latency introduce challenges for responsive data narratives at three stages of the visualisation architecture, and previously discussed issues of multi-device compatibility stifle the scope of

[6] specific frames or instances where the user can manipulate the interface (see Vande Moere et al., 2012)

interaction modalities (Childs et al., 2013). To remedy this disparity of compatibility in interaction models across the myriad devices that consumers use for everyday computing, first steps have been taken towards establishing a taxonomy of interaction terms that are independent of device and technology, and that build on low-level user events and signals to form a declarative language that aims to simplify interactivity design in visualisation and beyond (Satyanarayan, Wongsuphasawat & Heer, 2014). *Brushing*, *Panning* and *Linking* are three useful verbs from this taxonomy that offer semantic and intuitive language for describing visualisation interaction.

Ultimately, interactivity in narrative data visualisation aims to activate agency in the user by engaging them simultaneously with visualisation as entertainment, and data as a source of empowerment, as discussed in the first chapter of this paper. User agency in data narratives translates as the ability of the user to navigate their own trajectory of interest, insight and understanding through the elements of the database depicted in the visualisation. Finding alternate explanations to data trends or factually verifying personal theories through visualisation are examples of agency given to users who engage with casual data visualisations. Real instances of user agency will be identified in the third chapter as they correspond to the interaction points of the chosen casual data visualisation applications.

2.2.6 Social Collaboration

Using visualisations as a locus for community discussion and understanding is an emerging and underexplored dimension of data visualisation, especially as it becomes more accessible for casual users (Heer, Viégas & Wattenberg, 2009; Isenberg et al., 2011). As discussed in Chapter 1, Sprague & Tory (2015) identify social interaction as a key motivating factor for users of casual visualisations. Platform, structure and interactivity in narrative data visualisation concern the use of visualisations as collaborative artefacts, where social spaces located around a particular visualisation become forums of shared insights and debate among users with different areas of expertise. Reconfiguring standard processes of communication, sensemaking, and even design itself to co-located and distributed processes is an affordance of collaborative visualisation and an emerging trend that supports the identified model of fluid interaction desirable in modern computing (Isenberg et al., 2011). Functional goals of collaborative visualisation focus on the ability to preserve, revisit, save and share visualisation states as well as user insights (Segel & Heer, 2010), and researchers studying responses to collaborative visualisation platforms have observed casual and fun interpretations of data alongside deep analyses and hypotheses (Wattenberg & Kriss, 2006). Collaboration has the potential to enrich data visualisation narratives, and carefully considered integration into existing social spaces online allows individual narratives to interact with the vast and infinite hypertextual narratives of contemporary digital society. Creators of data visualisations must recognise social collaboration as an underexplored facet of the medium. Embedding engagements with data visualisation into everyday human-computer interaction strengthens the identified function of data visualisations as narrative artefacts; means of telling stories about the human experience.

Designing for social collaboration places the user and their needs at the centre of the visualisation functionality, and so collaboration design practices may crossover into many best-practice approaches for narrative design in data visualisation. The inherent desire in users to share data views and perspectives among friends and peers where insight has been salient is recognised as one existing trope that collaborative visualisations can exploit to create interesting narratives (Viegas et al., 2004). Considering the range of user goals in engaging with a visualisation, and using this as a basis from which to enhance the user experience inclusively for many use cases including collaboration is an approach to narrative media design that can be adopted for this emerging medium. Heer, Viégas & Wattenberg (2007) identify and prototype features necessary to enable asynchronous collaboration of users in data visualisation. *Application Bookmarks,* which allow the user to revisit and share frames of a visualisation; and *Embedded Discussion*, where there is a bidirectional link between the thread of conversation and the area of the visualisation that is being discussed, are noted as catalysts of collective analysis (ibid.). There is a potential for data visualisation to adopt familiar social design features from mainstream digital media platforms such for music and video. SoundCloud and YouTube, two popular streaming sites, have implemented features that enable embedded discussion and application bookmarking by using timestamp hyperlinks in comment threads, sometimes explicitly and sometimes as implicit functions. Devising a generic spatiotemporal location language for dynamic data visualisations to use as hyperlink references on social media may be a source of important and progressive research in the area of integrated and collaborative narrative data visualisations.

Mackinlay (2009) argues that interesting stories emerge where society interacts with data and, through these social stories, there is collective understanding. For example, the financial data that told of the downfall of the global banking industry may have been better processed, presented and understood if it existed within a space where collaborative interaction and asynchronous public conversation was enabled (ibid.). In the third chapter of this paper, a key goal of the analysis is to assess the chosen applications of casual data visualisation in terms of their existing and potential collaborative functionality.

2.3 Chapter Conclusions

Data visualisations are narrative artefacts. The goal of this chapter was to examine perspectives on data visualisation as narrative artefacts by virtue of their emerging use for casual data interactions. Repositioning the technical field of visualisation within the realm of liberal arts allows for new thinking and new ideas to be generated; casual data visualisation is an evolving area that requires such ideas.

This chapter first defined narrative and explored its pertinence to the database in the current digital media landscape. From this, the findings of Chapter 1 were incorporated into an analysis of dimensions of narrative data visualisation with an emphasis on casual use cases. Throughout the chapter, useful language and terminology have been noted for use in later analysis, and the discovery of helpful and generic data visualisation language for casual users was an intended outcome of this chapter.

Many observations have been made about how the narrative medium of data visualisation can be addressed by designers and end users. Data visualisations appear across a range of platforms and can be classified by their genre. Designers of data visualisations must be aware of the areas of design that affect narrative success, identified in this research as platform, structure, stylistic embellishment, interactivity and social collaboration. Considering the implications of each of these dimensions can better inform the visualisation designer as to how their work will be consumed and explored by the user; using the terminology highlighted throughout the chapter presents a framework through which better communication between designers and consumers of casual data visualisations can be generated. The third chapter tests the use of this framework.

3 Application of Framework to Casual Data Visualisation Systems

In this chapter, the framework of narrative terminology devised in chapter two will be applied to three existing applications of casual data visualisations. The systems are:

- **Rap Stats:** A web-based n-gram viewer built on a comprehensive cultural database of rap music lyrics dating from 1988 to the present day. The application allows the user to look up and compare the popularity of a given word in the history of recorded rap music (Genius, 2015)
- **Reporter:** An iOS lifelogging app that uses notifications and user-defined questions to record explicit and ambient data in order to gain otherwise unknowable insight into aspects of the user's life (Felton, 2015b)
- **YouTube Trends:** A set of visualisation tools including a map, dashboard and blog for exploring and comparing currently popular videos on YouTube, displayable per region across different demographics (YouTube.com, 2015)

3.1 Motivation for Selected Systems

The goal of this analysis chapter is to apply the devised framework to the selected systems in order to better understand their position within the vast and expanding field of casual data visualisation. Three casual data visualisation systems have been selected to qualitatively investigate the analytic functionality of the defined narrative framework. Though a limited set of applications, each system is considered by this author to be representative of an exciting area of casual data visualisation that is ripe for advancement and investigation in future research and real-world visualisation development. Rap Stats represents the slew of simple interface tools available for exploring online corpora; Reporter is an example of a personal data visualisation system for lifelogging; YouTube Trends is a casual application that attempts to digest the millions of data points that describe cultural interactions with social media in a sensible, logical interface. The three systems are justified as casual data visualisation applications suitable for analysis for this research as they exist as products of the identified era of Big Data examined in the first chapter; none of these applications were feasible prior to the technological, sociological and economic changes that have allowed casual users to engage with personal and cultural data narratives as part of their everyday interactions with the world. The hypothesis of this analysis is that the narrative strengths and weaknesses of each system may be discovered through the application of the devised framework.

3.2 Analysis

Table 2 shows a summary of the findings of the analysis. The three casual data visualisation systems were analysed by applying the framework systematically to each and then qualitatively assessing the extent to which the system promotes or inhibits a particular dimension of narrative. Through this systematic process, narrative strengths and weaknesses of each casual data visualisation application were noted. The following section details the findings under the six dimensions of analysis.

Dimension	Rap Stats	Reporter	YouTube Trends
Genre	Single Frame Chart	Slide Show	Map + Chart + Magazine-Style
Structure	Reader-Driven	Martini Glass Structure	Drill-Down Story
Platform	Web Not optimised for Mobile	iOS Native Application	Web Not optimised for Mobile
Stylistic Embellishment	Very Little Prioritises Data Clarity Potentially Demotivating	Modern Design Considerations Minimal Colour Schemes	Almost None Potentially Demotivating
Interactivity	Single Point of Interaction Limited Agency	High Levels of Agency Aesthetic Manipulation Functional Manipulation	High Levels of Agency Functional Manipulation Detailed Data Manipulation
Social Collaboration	Social Media Integration Chart Image Download	Only Possible by Exporting Data	None

Table 2 – A Summary of Findings from the Application of the Devised Framework of Analysis to Three Selected Casual Data Visualisation Systems

3.2.1 Genre

Rap Stats is a simple chart. There is only one frame per visualisation. The displayed chart is not annotated beyond including tooltips with specific data at each point, and it is not built into a slide show or magazine style presentation. Figure 1 shows the Rap Stats data visualisation interface.

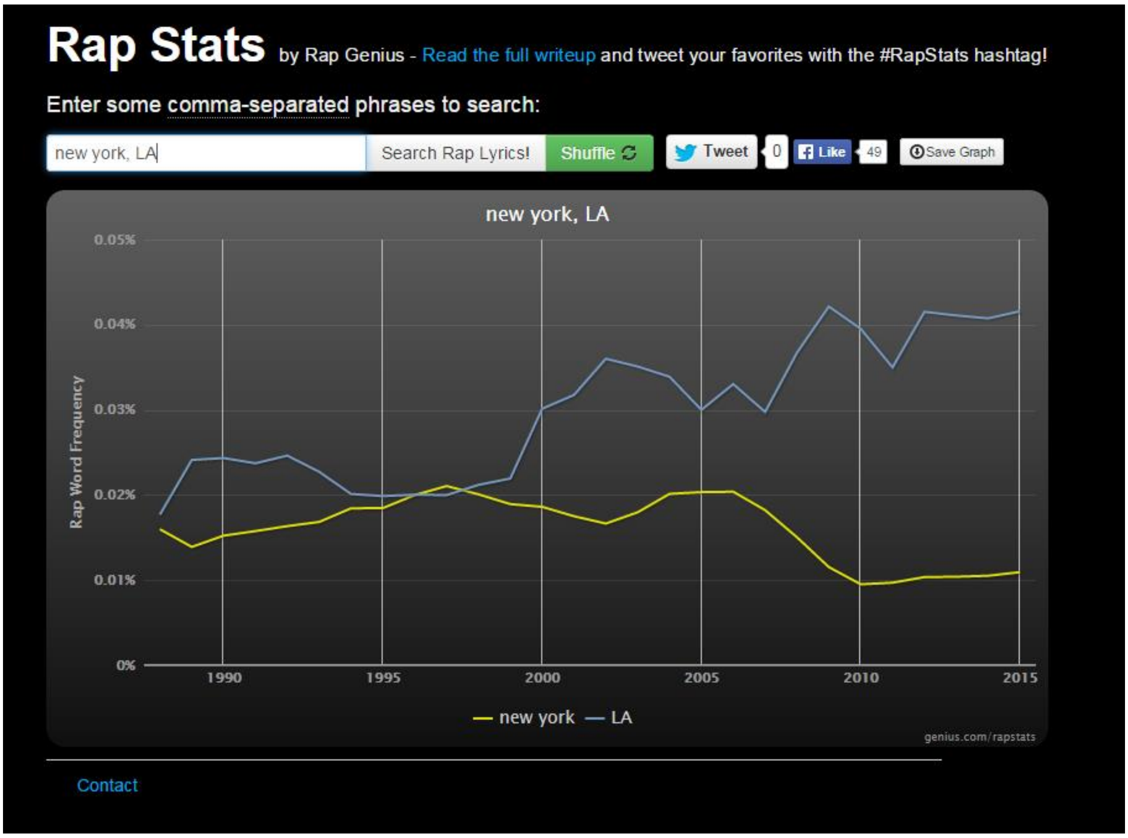

Figure 1 - The *Rap Stats* Interface (Genius, 2015)

Reporter falls within the slide show genre, and there are various types of charts, diagrams and tag clouds displayed in each frame. There is no chart annotation or supporting text of any kind in the visualisation. Figure 2 on the next page shows sample frames from the Reporter app.

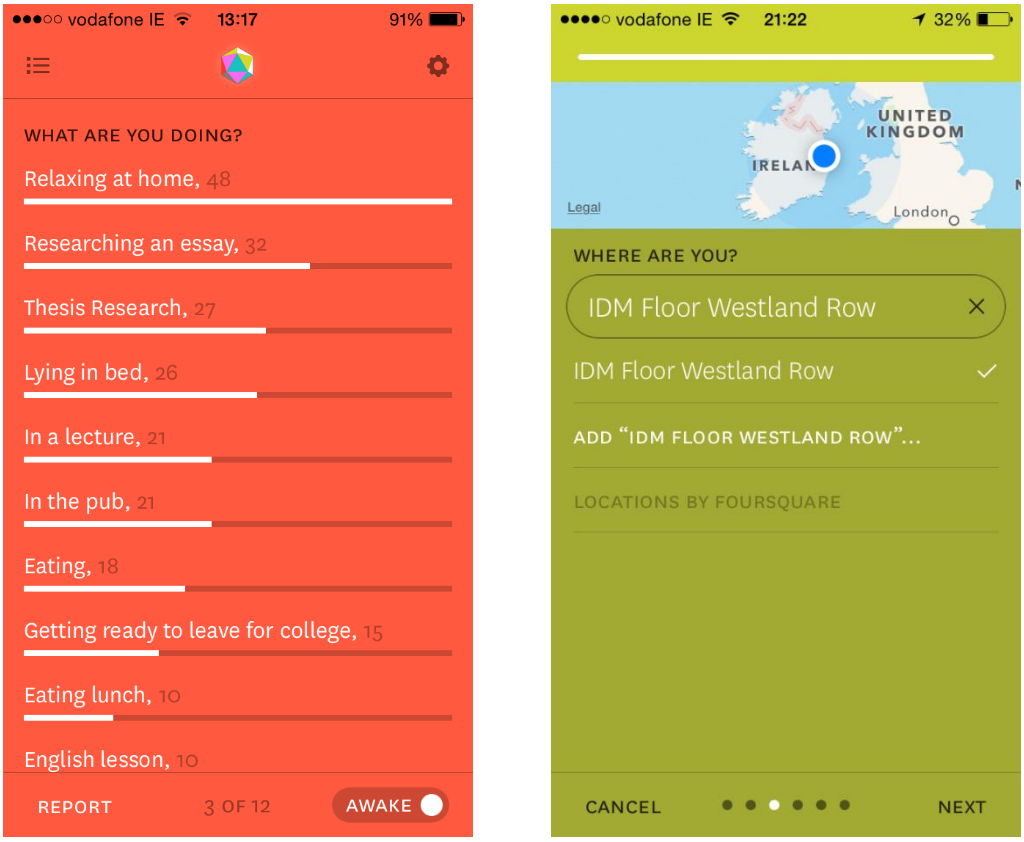

Figure 2 – Sample Frames from the *Reporter* App (Felton, 2015b)

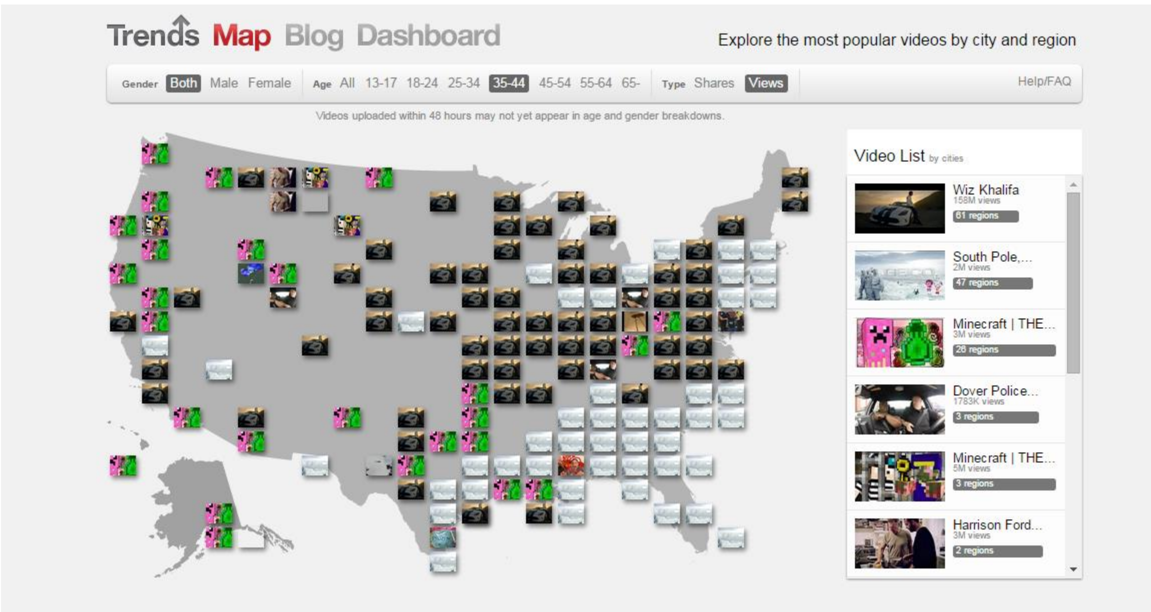

Figure 3 – The *YouTube Trends* Interface (YouTube.com, 2015)

YouTube Trends combines map, chart and magazine style genres to present an array of tools for examining the same dataset. There are multiple chart types and the application includes a lot of supporting text in the magazine style layout that offers context to the data and narrative, while annotation is limited to descriptive details about the videos. Figure 3 shows the map frame of the YouTube Trends data visualisation interface.

Assigning each system to one or more of the identified genres in the framework was straightforward, but the process of finding a ‘best fit’ for each system was inaccurate. Further delineation of data visualisation genres with an emphasis on emerging casual systems could perhaps uncover more accurate instruments of analysis for this dimension.

3.2.2 Structure

Rap Stats is a reader-driven narrative. There are suggestions by the author as to what kind of search terms give interesting results and there is a ‘shuffle’ button to aid the user if they are unable to think of their own search terms. However, the narrative outcomes of Rap Stats rely heavily on the user being creative with their investigation of the database through the system. There are opportunities to compare data using the visualisation, but spatial proximity of data points and level of detail are fixed. Logical flow and order are not relevant to this single-frame visualisation. However, a clear introduction to the context of the data has been overlooked by the system, where it may have been helpful for engaging the user in the narrative.

Reporter aligns with the Martini Glass Structure of data visualisation narratives. Upon first engaging with the app, the user is asked author-defined questions and the system displays charts of the recorded data. This gives the user context. A tutorial shows the user where to add and edit questions, and from this point the user is aware of the scope and logical flow of the system and so can be more autonomous in their use of the app.

The structure of YouTube Trends is a balance of author-driven and reader-driven, and the system can be classified as a Drill-Down Story, where difference level of detail and in-depth comparative analysis are available to the user. This structure allows data trends to be explored, revisited and recontextualised by confident users, while users unfamiliar with the system can still gain insight into popular YouTube videos through the initial visualisation state.

The analytic framework deals with structure successfully. By locating a system along the spectrum of reader-driven to author-driven, the basic narrative structure of the visualisation is easily observed. Additionally, the identified structures of Martini Glass and Drill-Down Story are helpful for approaching a semantic and qualitative classification of the systems.

3.2.3 Platform

Rap Stats is a web-based application. The combined search bar and n-gram chart that makes up this visualisation platform are evocative of the Google Trends application. This familiar and accessible platform choice foregrounds the corpus of texts in the database, rather than innovative design and interaction, as the primary vehicle of narrative. The system is usable on mobile but is not optimised or reconfigured for small devices or touch navigation. This results in the visualisation appearing differently across device models and sometimes responding unexpectedly to user input.

Reporter is a mobile application for iOS. As a native app, the system has access to sensors and data sources provided by the iPhone platform including ambient noise level through the microphone, accurate GPS and altitude data from the cellular network, and pedometer data from the accelerometer. A web application could not collect or interpret the same data easily. Reporter exploits the platform as a data source, and the ambient data drives the possible narratives of the visualisation system. Furthermore, Reporter makes use of random notifications to prompt user input, another affordance of the always-on mobile platform. However, the charts displayed in the app do not take advantage of the touchable interface of the iOS platform. The charts are viewable by the user but there is little interaction at the visualisation stage, though JSON and CSV files of raw data are available to the user at any point. In this way, Reporter succeeds in shielding the user from complexity while affording them the opportunity to deeply explore their data and expand the narrative.

YouTube Trends is a web visualisation. There is no dedicated mobile version of the application, so navigating the interface using touch is challenging. The application uses hover action throughout the system to display details to the user and this design choice proves restrictive on mobile platforms. Only users accessing the system from a point-and-click GUI are afforded the ability to quickly and easily explore and compare the data. The web platform allows the visualisation to stay updated with the latest YouTube viewer figures, but there is a significant flaw in the platform regarding the promptness at which the system updates. A disclaimer notifies the user that videos uploaded to YouTube within the last 48 hours may not be represented accurately in the visualisation. This affects the possibility to keep updated with emerging viral video trends, and so the software and network latency in the platform become a barrier to possible narratives communicated through the system.

Analysing the hardware, software and network considerations of a system under the common dimension of platform is certainly helpful for easily gaining a preliminary understanding of the narrative implications of each technology, but each area is itself a broad dimension of analysis. For the purpose of narrative design, then, the dimension of platform is a useful analytic instrument. However, the framework does not satisfy specialist analysis of casual visualisation systems.

3.2.4 Stylistic Embellishment

The design considerations for Rap Stats prioritise data clarity, simplicity and user-driven discovery. There is little stylistic embellishment. The visualisation has a high proportion of data-ink, with minimal axis labelling and sparse grid markers. Imagery is not used in the data visualisation; the data itself is foregrounded through the design. The entire interface is made up of the defined chart area, the text input area and the data points. Primary and secondary colours are used to represent data points, while the axes, backdrop and grid lines are monochromatic. The design choices for this visualisation give the charts a traditional and scientific aesthetic, which presents an interesting juxtaposition to the niche cultural nature of the dataset. The absence of bright colours and enticing images may be a demotivating factor for some casual users, but the simplicity of the design reflects the simplicity of the system itself, and enables easy access to user-driven narratives through the system.

Reporter has a high proportion of data-ink. Almost the entire screen space is dedicated to chart display, and so data clarity is central to the design. The charts in Reporter are standard statistical charts, but the bright colours and modern user interface design approach give the visualisation a contemporary aesthetic. This is potentially motivating for the user, and the design does not distract from or obscure the data. The user can select from five colour schemes. Shades and tints of a single colour rather than discrete colours are used to distinguish user interface features within each colour scheme. There are no images or animations used in the Reporter app. Overall, the design augments the standard display of data charts in a slide show with clever use of colour and layout; the potential for user-driven narrative is positively influenced by the stylistic embellishment.

YouTube Trends makes economic use of data-ink by using a small thumbnail of each video as data points over a basic map of the United States. Below the map, a bar graph shows the overall percentage of views each popular video has and there is no unnecessary text, imagery or user interface components. Exploring the visualisation does not require any familiarity with modern web interfaces and there are no animated transitions to distract from the data; the design does not restrict access to interesting narrative. Like Rap Stats, the interface of YouTube trends is a potentially demotivating aspect of the design as it is mostly grey with little embellishment.

Like platform, stylistic embellishment is a broad dimension that is useful for quickly gaining a general understanding the narrative implications of a system’s aesthetic. Identifying potential distractions from the core data is made easy by applying the framework, but certainly more foundational knowledge of whether casual visualisation systems benefit from embellishment is needed.

3.2.5 Interactivity

User agency with Rap Stats is relatively low, relative to the size of the database that the system is built on. The user can query up to eight terms to be displayed on the chart, and there are *and* and *or* operators available for more accurate results. However, the visualisation itself cannot be manipulated by the user and there are no options to configure, compare or contextualise the data beyond the basic line graph. The only interaction point is the search bar. Temporal, spatial and aesthetic parameters of the data narrative are set, and agency is limited to the search queries the user provides.

In addition to basic interactivity functions such as swiping through the slideshow by, Reporter offers the user novel agency in two key ways. The aesthetics of the visualisation can be configured by the user, though this is limited to a set array of colour schemes and chart displays. However, even this limited agency over the design enables the user to engage more personally with the visualisation narrative. In addition to aesthetic customisation, the application allows the user to define their own questions. This functional interactivity has the potential for each user to experience a personal and unique narrative through the casual data visualisation system. In terms of interactive engagement with the data at the visualisation stage, interactivity is low. Linking datasets to generate truly unique insights is only available to the user if they export and visualise the raw data externally.

Through its Drill-Down Story structure, YouTube Trends affords the user a lot of functional agency in comparing and linking visualised data. Hovering over thumbnails and selecting different tabs allow the user to manipulate the interface, and there are several points of interaction in each view. Using these interaction points, the user can focus on data from different regions and demographics at their own pace. The narrative scope afforded to the user by the interactivity in this application is large, but the aforementioned data latency is a barrier to a fully-realised interactive narrative.

By applying this framework, valuable insight into the positive and negative implications of interaction points in a visualisation is gained. The influence of interactivity on narrative becomes evident to the designer, and additional insight into the dimensions of structure and platform are enabled through analysing interactivity in this way.

3.2.6 Social Collaboration

The Rap Stats visualisation can become the locus of data-driven conversations in several ways. Interesting visualisations can be shared on Facebook and Twitter feeds and basic chart images can be saved for use in any number of ways by casual users. Though the visualisation system does not implement any innovative collaborative techniques, its familiar features of social media integration enable the narrative to extend beyond the application and become a catalyst of collective analysis.

As a lifelogging system, Reporter does not natively feature any means to share data stories. The app records personal data such as location and diet, and some users may use it to keep records of sensitive information about their relationships and wellbeing. A user who wishes to share their data story must export the raw data and present it using their own form of visualisation. In this way, Reporter does not restrict collaborative analysis, but it is not prioritised as a narrative outcome of the system.

YouTube Trends does not include the features identified to enable successful social collaboration. The narrative created through interacting with the data is between the individual user and the system; there is no way to share or comment on insights or data points. The initial state of the data visualisation is the same for all users at all times no matter where in the world they are accessing it from, but deeper exploration of the system enables more unique and personal narratives which could be the genesis of collaborative analyses if the functionality was enabled in the system. There is a comment section under each post in the blog tab, but the discussion encouraged here is not related to the data as much as it is the video being discussed.

Evaluating the ability for a data visualisation to be used socially is made simple by this framework. Analysis by this dimension involved identifying features within a system that extend the narrative beyond what is presented on the visualisation interface. Under this framework, opportunities for social collaboration are the most clearly observable features of any dimension of analysis.

3.3 Chapter Conclusions

The framework of analysis was applied to each system. Several strengths and weaknesses relating to narrative success were uncovered during the analysis that may have been overlooked if an empirical methodology was employed. For example, Rap Stats is a reader-driven application with a restricted interface, but it has the potential to tell interesting data stories through the use of social media. Many features of the app are limited, but these limits are justified by the multiple opportunities for social collaboration. Therefore, Rap Stats could be understood as a tool to support other narratives with factual data, rather than a self-contained narrative data visualisation. Conversely, Reporter and YouTube Trends allow a more detailed investigation of data within the system by a single user. They achieve this by giving the user agency over application functionality at the data input stage and the visualisation stage, respectively. These two systems are more complete data stories, but could further extend narrative possibility through social media integration.

Each dimension of analysis in the framework was found to influence the possible narratives the system could communicate, and this demonstrated that the taxonomy of terms is a useful tool for understanding casual data visualisations as narrative objects. However, some factors are not addressed by this framework.

Firstly, the understanding of casual data visualisation does not distinguish between visualisation design tools and end user systems. This makes it difficult to qualify the narrative dimensions of the system under the devised framework. The line between creative and consumer visualisation software is becoming increasingly blurred in the era of casual data, and a better understanding of the classification of visualisations along this spectrum would add much-needed context to the exploration of casual data visualisation as narrative.

In addition, the emergent state of casual data visualisation means that much is still unknown about real-world user engagement with these systems. Data visualisation and interaction design enthusiasts are the early adopters of the kinds of systems selected for this analysis, and many features of the applications would likely undergo a full redesign is the products were aimed at a mass market. However, the opportunities and barriers for narrative success identified in each casual data visualisation system provide a guideline for approaching such a redesign.

This application chapter concludes that the framework successfully enables insight into the narrative dimensions of casual data visualisation systems. However, refined and deeper knowledge of each dimension is necessary for engaging in low-level technical analysis of a particular aspect of a system, but the framework provides a common language for communication between disciplines at a high-level.

Conclusions

Casual data visualisation is an emerging area of interest both inside and outside academia. As a subset of data visualisation and data analytics; casual data visualisation spans the fields of computer science, art and design. The notion of casual data as a resource for non-expert users is understood to be a product of the Big Data era of modern computing. Casual data visualisation allows these non-expert users to engage meaningfully with the quantified data that powers the modern world. Understanding how casual user engagement with data can enable insight and entertainment is the concern of casual data visualisation research.

Simultaneously, narrative data visualisation has been the focus of a large amount of research in academia, and many papers have investigated the ability for data visualisation to be used as storytelling media. Narrative visualisation is another subset of the wider field of visualisation research. Understanding the motivations for users to engage with a visualisation as well as identifying potential inhibitors of use are two key factors for measuring the narrative success of a particular data visualisation.

This research paper has sought to integrate both subsets of data visualisation research into an examination of casual data visualisation as narrative media. By incorporating perspectives of narrative theory into an analysis of existing research into the narrative properties of data visualisation, a framework for analysing casual data visualisations as narrative media was devised. The framework consists of six dimensions of analysis; *Genre*, *Structure*, *Platform*, *Stylistic Embellishment, Interactivity* and *Social Collaboration.*

Applying the framework to three real-world applications of casual data visualisation demonstrated that it is a useful approach for investigating the narrative strengths and weaknesses of casual data visualisation systems. As the area of casual data visualisation expands and incorporates increasingly multidisciplinary practice, casual users as well as experts from discrete fields will need a high-level common language for visualisation design and analysis. This framework presents one such language. Using the analytic framework to consider casual data visualisations as multidimensional narrative media is a vital step in positioning these systems among every day consumer media, and simultaneously promoting interactions with data as salient narrative sources.

The changing nature of data visualisation demands new methodologies, and this paper offers a valuable taxonomy to researchers looking to explore casual systems. Applying my framework systematically across a larger corpus of casual data visualisation systems would yield valuable quantitative data with which to compare the qualitative findings of this paper, and I recommend this as a logical continuation of the presented work. Additionally, social collaboration in casual data visualisation is recommended as an area that is ripe for future study, and my analytic framework allows researchers to easily uncover narrative opportunities to this end.

On a final note, I would like to offer reflective criticism on my own methodology. As I progressed through the research period, my own understanding of casual data visualisation systems was consistently changed by the discovery of new visual applications of fun, cultural and personal datasets. Engaging with such novel applications of casual data was the most interesting area of this research, and a rigid classification of these diverse casual systems would have better informed my taxonomy. In light of this, I propose further research into the classification of casual data visualisation systems 'in the wild', and the subsequent reassessment of this paper's taxonomy to accommodate emerging systems.

Bibliography

Abbott, H. (2002). *The Cambridge introduction to narrative*. Cambridge, UK: Cambridge University Press.

Amberber, E. (2015). *15 Indian Big Data companies to watch out for in 2015*. [online] YourStory. Available at: http://yourstory.com/2015/02/indian-big-data-companies-startups/ [Accessed 28 Feb. 2015].

Bal, M. (1997). *Narratology* [Preface]. Toronto: University of Toronto Press.

Barbry, E. (2012). The Internet of Things, Legal Aspects: What Will Change (Everything). *Communications and Strategies*, [online] 87, pp.83-100. Available at: http://www.idate.org/en/Digiworld/Communications-Strategies/Archives/Archives_50_.html [Accessed 24 Mar. 2015].

Barthes, R. & Heath, S. (1987). *Image, music, text*. London: Fontana Press.

Bass, R. (1999). Story and Archive in the Twenty-First Century. *College English*, 61(6), p.659.

Bateman, S., Mandryk, R., Gutwin, C., Genest, A., McDine, D. & Brooks, C. (2010). Useful Junk? The Effects of Visual Embellishment on Comprehension and Memorability of Charts. *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*.
CACM Staff, (2014). Visualizations make big data meaningful. *Commun. ACM*, 57(6), pp.19-21.

Cafaro, F. (2012). Using embodied allegories to design gesture suites for human-data interaction. *Proceedings of the 2012 ACM Conference on Ubiquitous Computing - UbiComp '12*.

Chatman, S. (1978). *Story and discourse*. Ithaca, N.Y.: Cornell University Press.

Childs, H., Geveci, B., Schroeder, W., Meredith, J., Moreland, K., Sewell, C., Kuhlen, T. & Bethel, E. (2013). Research Challenges for Visualization Software. *Computer*, 46(5), pp.34-42.

Cleveland, W. (1994). *The elements of graphing data*. Murray Hill, N.J.: AT & T Bell Laboratories.

Cukier, K. & Mayer-Schoenberger, V. (2013). The Rise of Big Data. *Foreign Affairs*, 92(3), pp.27-40.

Dictionary.com, (2015). *Story*. [online] Available at: http://dictionary.reference.com/browse/story?s=t [Accessed 8 Apr. 2015].

Dourish, P. (2001). *Where the action is*. Cambridge, Mass.: MIT Press.

Eaton, C., DeRoos, D., Deutsch, T., Lapis, G. & Zikopoulos, P. (2012). *Understanding Big Data*. New York: McGraw-Hill.

Ehrenberg, A. (1975). *Data reduction*. London: Wiley.

Eisenstein, S. & Leyda, J. (1949). *Film form*. New York: Harcourt, Brace.

Elmquist, N. (2011). Embodied Human-Data Interaction. *ACM CHI 2011 Workshop "Embodied Interaction: Theory and Practice in HCI*, pp.104-107.

Elmqvist, N., Vande Moere, A., Jetter, H., Cernea, D., Reiterer, H. & Jankun-Kelly, T. (2011). Fluid Interaction for Information Visualization. *Information Visualization - Special issue on State of the Field and New Research Directions*, 10(4), pp.327-340.

Felton, N. (2015a). *Nicholas Felton | Feltron.com*. [online] Feltron.com. Available at: http://feltron.com/ [Accessed 1 Mar. 2015].

Felton, N. (2015b). *Reporter for iPhone*. [online] Reporter-app.com. Available at: http://www.reporter-app.com/ [Accessed 30 Apr. 2015].

Figueiras, A. (2013). A Typology for Data Visualization on the Web. *2013 17th International Conference on Information Visualisation*.

Freedman, J., Hayles, N., McGann, J., McGill, M., Stallybrass, P. & Folsom, E. (2007). Responses to Ed Folsom's 'Database as Genre: The Epic Transformation of Archives'. *PMLA*, 122(5), pp.1603-1608.

Frigo, A. (2015). *Website of the records generated for 36 years*. [online] 2004-2040.com. Available at: http://2004-2040.com/ [Accessed 1 Mar. 2015].

Genius, (2015). *Rap Stats | Genius*. [online] Available at: http://rapgenius.com/rapstats [Accessed 30 Apr. 2015].

Gershon, N. & Page, W. (2001). What storytelling can do for information visualization. *Commun. ACM*, 44(8), pp.31-37.

Giachetti, C. & Marchi, G. (2010). 'Evolution of firms' product strategy over the life cycle of technology-based industries: A case study of the global mobile phone industry, 1980-2009', *Business History*, 52, 7, pp. 1123-1150, Business Source Complete, EBSCO*host*, [Accessed 1 Mar. 2015].

Heer, J., Viégas, F. & Wattenberg, M. (2009). Voyagers and voyeurs. *Commun. ACM*, 52(1), p.87.

Hullman, J., Drucker, S., Riche, N., Bongshin Lee, Fisher, D. & Adar, E. (2013). A Deeper Understanding of Sequence in Narrative Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12), pp.2406-2415.

Inbar, O., Tractinsky, N. & Meyer, J. (2007). Minimalism in information visualization. *Proceedings of the 14th European conference on Cognitive ergonomics invent! explore! - ECCE '07*.

Information Management, (2015). *Top 10 Pure Big Data Companies*. [online] Available at: http://www.information-management.com/gallery/Big-Data-Pure-Play-Revenues-Companies-Ranked-10026249-1.html [Accessed 28 Feb. 2015].

Isenberg, P., Elmqvist, N., Scholtz, J., Cenrea, D., Ma, K. & Hagen, H. (2011). Collaborative visualization: definition, challenges, and research agenda. *Information Visualization - Special issue on State of the Field and New Research Directions*, 10(4), pp.310-326.

Kiousis, S. (2002). Interactivity: A Concept Explication. *New Media & Society*, 4(3), pp.355-383.

Kosara, R. & Mackinlay, J. (2013). Storytelling: The Next Step for Visualization. *Computer*, 46(5), pp.44-50.

Kosciejew, M. (2013). The Era of Big Data. *Feliciter*, 59(4), pp.52-55.

Klien, N. (2007). Waiting for the World to Explode. In: V. Vensa, ed., *Database Aesthetics: Art in the Age of Information Overflow*, 1st ed. Minneapolis: University of Minnesota Press.

Mackinlay, J. (2009). Technical perspective Finding and telling stories with data. *Commun. ACM*, 52(1), p.86.

Madhavan, J., Balakrishnan, S., Brisbin, K., Gonzalez, H. & Gupta, N. (2012). Big Data Storytelling through Interactive Maps. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*.

Manovich, L. (2001). *The Language of New Media*. Cambridge, Mass.: MIT Press.

Manovich, L. (2007). Database as a Symbolic Form. In: V. Vensa, ed., *Database Aesthetics: Art in the Age of Information Overflow*, 1st ed. Minneapoli: University of Minnesota Press.

Manovich, L. (2011). What is Visualization? *Visual Studies*, 26(1), Routledge.

Mander, J. (1978). *Four arguments for the elimination of television*. New York: Morrow.

McLuhan, M. & Lapham, L. (1994). *Understanding media*. Massachusetts: MIT Press.

Merriam-webster.com, (2015). *Story*. [online] Available at: http://www.merriam-webster.com/dictionary/story [Accessed 8 Apr. 2015].

Moore, D. (1998). Statistics among the Liberal Arts. *Journal of the American Statistical Association*, 93(444), p.1253.

Mortier, R., Haddadi, H., Henderson, T., McAuley, D. & Crowcroft, J. (2014). Human-Data Interaction: The Human Face of the Data-Driven Society. *SSRN Journal*.

Pfannkuch, M., Regan, M. & Wild, C. (2010). Telling Data Stories: Essential Dialogues for Comparative Reasoning. *Journal of Statistics Education*, 18(1), pp.1-38.

Pousman, Z., Stasko, J. & Mateas, M. (2007). Casual Information Visualization: Depictions of Data in Everyday Life. *IEEE Transactions on Visualization and Computer Graphics*, 13(6), pp.1145-1152.

Roberts, J., Ritsos, P., Badam, S., Brodbeck, D., Kennedy, J. & Elmqvist, N. (2014). Visualization beyond the Desktop--the Next Big Thing. *IEEE Computer Graphics and Applications*, 34(6), pp.26-34.

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), pp.1325-1332.

Rock, M. (1992). *Since When Did USA Today Become Our National Design Ideal?*. [online] 2x4.org. Available at: http://2x4.org/ideas/12/since-when-did-usa-today-become-our-national-design-ideal/ [Accessed 25 Apr. 2015].

Rosling, H. (2013). The Joy of Facts and Figures [Interview with Hans Rosling by Fiona Fleck]. *Bulletin of the World Health Organization*, 91(12), pp.904-905.
Saint, A. (2015). Where next for the Internet of Things?. *Engineering & Technology*, 10(1), pp.72-75.

Saraiya, P., North, C. and Duca, K. (2005). An Insight-Based Methodology for Evaluating Bioinformatics Visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4), pp.443-456.

Satyanarayan, A., Wongsuphasawat, K. and Heer, J. (2014). Declarative interaction design for data visualization. *Proceedings of the 27th annual ACM symposium on User interface software and technology - UIST '14*.

Scanlan, C. (2014). *What is Narrative, Anyway?*. [online] Poynter.org. Available at: http://www.poynter.org/news/media-innovation/16324/what-is-narrative-anyway/ [Accessed 7 Apr. 2015].

Segel, E. and Heer, J. (2010). Narrative Visualization: Telling Stories with Data. *IEEE Transactions on Visualization and Computer Graphics*, 16(6), pp.1139-1148.

Sprague, D. and Tory, M. (2015). Exploring how and why people use visualizations in casual contexts: modeling user goals and regulated motivations. *Information Visualisation*, 11(2), pp.106-123.

The Economist, (2010). *Data, data everywhere*. [online] Available at: http://www.economist.com/node/15557443 [Accessed 28 Feb. 2015].

Tufte, E. (1986). *The visual display of quantitative information*. Cheshire, Conn.: Graphics Press.

Tufte, E. (2006). *Beautiful evidence*. Cheshire, Conn.: Graphics Press.

Tukey, J. (1977). *Exploratory data analysis*. Reading, Mass.: Addison-Wesley Pub. Co.

Vande Moere, A., Tomitsch, M., Wimmer, C., Christoph, B. & Grechenig, T. (2012). Evaluating the Effect of Style in Information Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 18(12), pp.2739-2748.

Viegas, F., Boyd, D., Nguyen, D., Potter, J. & Donath, J. (2004). Digital artifacts for remembering and storytelling: posthistory and social network fragments. *Proceedings of the 37th Annual Hawaii International Conference on System Sciences, 2004.*

Viégas, F. & Wattenberg, M. (2007). Artistic Data Visualization: Beyond Visual Analytics. *Online Communities and Social Computing*, pp.182-191.

Wattenberg, M. & Kriss, J. (2006). Designing for social data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 12(4), pp.549-557.

Weinbren, G. (2007). Ocean, Database, Recut. In: V. Vensa, ed., *Database Aesthetics: Art in the Age of Information Overflow*, 1st ed. Minneapolis: University of Minnesota Press.

Wohlfart, M. & Hauser, H. (2007). Storytelling for presentation in volume visualization. *EUROVIS'07 Proceedings of the 9th Joint Eurographics / IEEE VGTC conference on Visualization*, pp.91-98.

Womack, R. (2014). *Outline of Principles from William Cleveland's Elements of Graphing Data*. [online] ryanwomack.com. Available at: http://ryanwomack.com/IASSIST/DataViz/Data/Cleveland.pdf [Accessed 24 Apr. 2015].

World Economic Forum, (2011). *Personal Data: The Emergence of a New Asset Class*. Geneva: WEF.

Young, N. (2012). *The Virtual Self*. Toronto, Ont.: McClelland & Stewart.

Youtube.com, (2015). *YouTube Trends Map - YouTube*. [online] Available at: https://www.youtube.com/trendsmap [Accessed 30 Apr. 2015].

Zacks, J., Levy, E., Tversky, B. & Schiano, D. (2002). Graphs in Print. In: M. Anderson, B. Meyer & P. Olivier, ed., *Diagrammatic Representation and Reasoning*, 1st ed. London: Springer-Verlag.

Zaslavsky, A., Perera, C. & Georgakopoulos, D. (2012). Sensing as a Service and Big Data.*Proceedings of the International Conference on Advances in Cloud Computing (ACC)*, (8), pp.21-29.

Tables & Figures