Playing with Complexity
An approach to exploratory data visualisation

Ben Ennis Butler, Doctoral Candidate, University of Canberra
Dr Sam Hinton, Assistant Professor in Digital Design, University of Canberra
Dr Mitchell Whitelaw, Associate Professor in Digital Design, University of Canberra

Introduction

In an era when 'information overload' is a cliché, and the verb 'to google' is in common use, data becomes cultural material. Networked information is woven ever more tightly into the fabric of contemporary society; data is the medium of our collective present as well as, increasingly, our history. As such it is vital that data also becomes a material for design. This means not only designing its surface or appearance, but working with its deep structures and attributes, designing its flows, interactions and representation to create new forms.

In our own practice we work with digital cultural collections. (Whitelaw, 2009; Hinton and Whitelaw, 2010; Ennis Butler 2010) These data sets offer a case study in data as a cultural material; in the collections of institutions such as the National Archives and the National Gallery, we literally find the traces of our culture in digital form. These collections, which are increasingly available online, also pose an interesting challenge in their size and complexity. Access to these collections is typically provided through search and browse interfaces, but as de Caro (2010) explains: 'the search function has proven an unsatisfactory solution, as it is only capable of locating a specific piece of information that the user can identify precisely.' Search-based interfaces demand knowledge of the collection; but often we can only discover a collection's contents through searching! In our work we aim to develop an alternative approach: rich visualisations of these collections that encourage open-ended exploration and allow the user to develop a richer and more complete understanding of the data.

At the heart of this process is the difficulty we frequently encounter when faced with complexity. Do we take an approach where we try to manage the complexity of the data before it is presented to an interested person (a gallery visitor, for example), or

do we embrace complexity and present it to the user warts and all, and allow them to explore it, to derive their own meanings and relationships from the data presented?

We aim to embrace complexity, and the challenge it poses. In this paper we situate our work in a wider theoretical context. We begin by examining the concept of complexity, before going on to look at visual information seeking as an interactive computer based approach to dealing with complexity in large data sets. The paper then goes on to look at the concept of play and ends by suggesting that playfulness is an important and under-explored aspect of interactive data visualisation. Playfulness may be a key feature of visualisations that embrace complexity.

Complexity

Complexity is an elusive concept, with a range of specialist meanings in different fields. Here we are concerned with a generalist or ‘everyday’ notion of the complex, which refers to systems with many elements, and many relationships between those elements. Weaver (1948) makes a useful distinction between ‘organised’ and ‘disorganised’ complexity. Disorganised complexity refers to systems with many independent or unrelated elements; their sum often resembles randomness or noise. More relevant here is the ‘organised’ complexity of a system with elements that are interrelated to varying degrees; the system as a whole is neither completely ordered (uniform) nor completely disordered (random), but in a richly structured in-between state. This state of organised complexity is ubiquitous: we find it in weather patterns, living organisms, and human society and culture.

The subjective experience of complexity has two sides. On one hand complexity is associated with effort or work: computational measures of complexity link it to the number of steps a required to complete a procedure. Similarly in everyday usage, a complex problem is more difficult to understand than a simple one. Approaches based on cognitive load theory often seek to minimise complexity, and thus mental effort. (see for example Hollender et al. 2010) On the other hand, complexity is also related to interest, engagement, and aesthetic pleasure. Most would agree that structures combining order (or unity) and diversity are more interesting than those that are completely ordered, or completely random. Drawing on evolutionary, psychological and aesthetic theory, Peter Smith goes further, arguing that ‘the underlying principle of aesthetic experience is that of complexity giving way to

orderliness.' (10) Smith contends that complexity is not rewarding in itself, but that the discovery of order within complexity - of Wheeler's underlying 'organisation' - is intrinsically pleasurable. Here Smith echoes the work of Ramachandran and Hirstein (1999), who argue that the perceptual process of discovering patterns or correlated features carries a hard-wired pleasurable reward. (21) In both cases, aesthetic pleasure is associated with discovery or the acquisition of knowledge; in fact knowledge - the discovery of new patterns within complex perceptual data - is the evolutionary reason for the existence of aesthetic pleasure. Like many other practitioners in data visualisation, we aim to develop visualisations that both promote discovery and insight, and provide an aesthetically rewarding experience. Smith shows that these two goals may in fact be one and the same.

Interactive Data Visualisation

Interactive data visualisations are one method of allowing humans to interact and derive meaning from complex data sets, such as cultural collections. Rodenbeck remarks that visualisation is emerging as a medium in its own right (2010) and Manovich sees visualisation-like approaches (which he terms 'cultural analytics') as a way of 'generating new approaches for studying cultural history and contemporary culture.' (14)

Some commentators have identified two emerging tendencies in data visualisation, distinguished by the intent of the visualisation's design. This divergence has recently become a source of some debate within the data visualisation community. (Few, 2011; Bertini 2011) In exploratory data visualisation the intent is to present the data 'as is' so that the viewer can make their own determinations about the meaning of the data. Here the designer attempts to maintain a neutral position with respect the meaning of the data set that is being visualised. In communicative visualisation the goal of the designer is to present a particular view, to show the user what the data means, or to use data visualisation as a persuasive medium. Many infographics that accompany newspaper articles fall into this category. In this paper, and in our work more generally, we focus on exploratory data visualisation.

Like many others in this field, we draw on the work of Ben Shneiderman, who described a Visual Information Seeking (VIS) approach to data visualisation in 1996. This approach describes how users engage [with data] when seeking information in

any interactive environment. This VIS approach is summarised as ‘Overview first, zoom and filter, then details on demand.’ In other words, start with all the data, then zoom and filter out unwanted content, before selecting an item to finding out more about it. As Keim remarks the intent is to present the data in a visual form that allows the user to gain insights, interact with, and draw conclusions from it. (2002)

As a method for exploring large data sets this approach has proven to be particularly effective (Keller and Tergan, 2005) as it provides the user with an immediate appreciation for the size and diversity of the collection. (Hornbæk and Hertzum, 2011) Interactivity in data visualisations is particularly important, as it provides the user with mechanisms for handling complexity (Oliveira and Levkowitz, 2003) and enables them to engage in visual information seeking as they filter and focus on parts of the data that they find interesting. This process, which will be different for each user, will lead to greater exploration and understandings of the data. (Keim, 2002)

Shneiderman’s VIS approach has been very influential in visualisation research; it is cited over 1400 times on Google Scholar. However a survey of 60 papers on the topic by Hornbæk and Hertzum found that while all the papers refer to the ‘overview’ as a key concept of a successful visualisation, none of them explicitly define its meaning. (2011) While Shneiderman uses the term in a more technical sense, Spence argues that visualisation is a human cognitive activity and has nothing to do with computers, thus an overview may not be obtained immediately but may require time and effort. (2007) This is consistent with Smith’s account of the aesthetics of complexity, where aesthetic pleasure is a hard-wired reward for the cognitive effort involved in discovering order within complexity.

Play

The concept of play seems to present a useful way of thinking about our approach to interactive visualisation of complex data. We develop our ideas about play below with reference to computer games, but emphasise that we do not see play and games as the same thing. We see play as a mode of experience, which encompasses more than games, and we are not advocating a simple ‘gamification’ of data visualisation. Instead, we understand computer games as examples of playful digital software that engage with complex data (‘game data’) to produce entertaining interactive

experiences. The question we ask, then, is without necessarily making data visualisation a game, how can elements of play and playfulness be usefully employed in data visualisation to aid users in deriving meaning from a dataset?

A reasonable place to begin is the problematic concept of play. Until recently play has been severely under-theorised as a cultural feature of human societies. It has attracted little attention from sociologists presumably because it was seen as separate, unproductive and possibly childish behaviour, therefore assumedly unworthy of significant scholarly attention. Over the last decade, as computer games have become recognised as an economically and culturally significant medium, scholars have again returned to the idea of play, especially as it relates to games.

In early work Huizinga (1950) and later Caillois (1957) both define play in general terms as a special kind of activity that occupies its own time and space and exists outside of regular experience. It is safe, separate, and pleasurable. Huizinga described this exceptional kind of experience as a magic circle, in which the rules of everyday life are temporarily suspended.

Caillois described two ends of a spectrum on which play exists, which provides a useful conceptual framework for considering the role of play in data visualisation. On one end of the spectrum, which he called *ludus*, is highly structured play where the player's behaviour is governed by a strict set of rules. At the other end of the spectrum is *paidia* - play that is characterised by the complete lack of rules. Different activities sit at different points within this spectrum. A game like chess is positioned closer to the ludus end of the spectrum, while children's make-believe role-play is towards the paidia end.

Following Caillois, recent games studies scholarship focuses on understanding the difference between play and games, in an attempt to more clearly demarcate the developing field of game studies. A number of scholars describe games as a class of play that is defined by an adherence to structure and rules. Juul (2003), for example, reviews various definitions of games, and comes to his own definition, describing six characteristics of 'gameness' that explicitly exclude what he calls free-form play. Play, then, is a superset of activities or behaviours of which games are only one kind. Games are seen as a special instance of play.

An alternative view is proposed by Malaby. (2007) Malaby offers a useful counter position in which he critiques what he sees as the exceptionalism in much game and play scholarship that continues to separate play from everyday activity. One of the critical points that Malaby makes here is the idea that play can be usefully thought of as a mode of experiencing reality, rather than a separate activity that has no bearing on the real world.

For Malaby games are a set of processes that are linked to experience; many non-game activities have playful elements, just as many games have strong relationships with the real world. Using examples of both digital and non-digital games, Malaby provides empirical evidence to show that play is not always separate, safe nor necessarily pleasurable, and that these notions of play are culturally nuanced and socially constructed. He also points to work by scholars like Sherry Ortner, who defines the distinction between work and play/leisure as a 'modernist affectation' that results from the 19th century construction of the idea of work. (8)

If play can be considered as a mode of experience that occupies the same space as the rest of experience, rather than as a walled off space defined by the magic circle, then the need to separate play from games is reduced, while at the same time a range of activities can be revealed as playful, or as having playful qualities. This opens the way to consider the role of play in many areas of life beyond entertainment, including education and human computer interaction.

Play and Visualisation

In data visualisation engaging with ideas of play may provide valuable tools for designers, particularly if we begin by thinking about the role of data and complexity in computer games. While many software applications deal with complexity that arises from large, complex digital data sets, games are centrally concerned with play and the effective engagement with complexity that facilitates that end.

In most computer applications, such as an income tax program, or even an online library catalogue, complexity is regarded as a problem that the user needs to be insulated from. Indeed, it is a major goal of software interface designers to simplify

and reduce complexity, to provide an interface that hides complexity by structuring and organising data. Here the goal of the software is to sort, analyse and filter to produce certainty. (Dibble, cited in Malaby)

By contrast, games deliberately generate unpredictability using random number generators and complex interactions between variables to create complexity that the player then has to deal with as a challenge that needs to be overcome. A good example of this is the Bell and Braben's 1984 game *Elite*. In *Elite* the player explores a vast galaxy of stars and planets, interacting with enemy spacecraft and buying and selling commodities in an effort to obtain 'elite' status. The universe created in this game is procedurally generated, with simple algorithms used for generating unique planet names and star system details. The entire game is essentially a way of visualising and exploring the complexity that results from a random number generator and some clever (but simple) generative algorithms. Meier's *Civilization* games use complex simulations involving tens of thousands of data points that the player controls to arrive at a desired outcome. Mastery of the game implies mastery of the complexity generated by the game.

Malaby refers to this kind of generative complexity as 'contrived contingency', and it forms a fundamental component of many computer games. Without this complexity the games would be predictable and unchallenging. In computer games, then, complexity, or more precisely, discovery or imposition of order within the complexity is something to be embraced, even enjoyed. Again, this echoes Smith's contention about the rewards of discovering order in complexity.

Given the importance of complexity in digital games, it's plausible to suggest that digital games are elaborate forms of interactive data visualisation. As we have seen, digital games are computer programs that, like exploratory visualisation, deal with uncertainty, complexity and contingency. This is particularly obvious in a game like *SimCity* or *Civilisation*, where Shneiderman's mantra of 'overview, zoom, filter then details on demand' is commonly employed to provide the player with a means to control and comprehend the millions of variables that underlie the game mechanic.

So, if games can be seen as kinds of data visualisation where complexity plays a central role, could the reverse also be true? Could interactive visualisations of complex data be playful? Clearly, interactive data visualisation deals with complexity,

although in place of the contrived complexity created by algorithms as seen in digital games, interactive visualisations engage with the complexity that extends from a data set generated by real world processes. Engaging viewers with the data visualisation in a playful way allows them to engage with the complexity of a real data set, to explore and discover links through a playful mode of experience.

This does not imply that playfulness in data visualisation necessitates the literal construction of some kind of data-based game ('gamification'), although this is one possible approach. (see, for example Diakoploulos et al. 2011) Instead we understand play in data visualisation as being closer to the paidia end of Caillois' spectrum, where interface elements encourage viewers to engage with the visualisation in a playful and exploratory way. Ideally the viewer is drawn in to the data by the interface, and encouraged to explore the unknown landscape that the visualisation lays out before them.

While we would argue that playful elements are present in many interfaces, some data visualisations more clearly illustrate these attributes. Harris and Kamvar's *We Feel Fine* is a good example of a data visualisation that engages the user through playful interface elements. In *We Feel Fine* the data set is presented as moving particles that respond to the mouse as the viewer moves around and clicks within the interface. This interface is engaging because it is fun to play with, and encourages deeper exploration that is heightened by a desire to learn more about the interface and the data it represents. Whitelaw's Visible Archive (2009) also has playful components, and encourages the viewer to explore the data set by playing with and combining word frequencies and constantly providing the user with feedback about the results of their explorations. Similar techniques are employed in Hinton and Whitelaw's Commons Explorer. (2010)

Conclusions

In this paper we have looked at the challenges of visualising large and complex digital collections. We have noted that digital cultural collections, like other information forms, are typically presented in a way that minimises their apparent complexity. Search and browse displays both flatten the complex structures of a

collection into a simple ordered list, showing only a tiny fraction of the collection at a time.

Our approach is inspired by Shneiderman's VIS and the 'show everything' motto of Stamen (Jones, 2009), in which scale and complexity are valuable assets that can be engaged with through playful interfaces. We conceive of play here not as something that is separate to everyday life, but as something that is embedded within it, a mode of experiencing reality. Suspicions that a playful mode of experience may be a useful approach to data visualisation is supported by closer examination of digital games, which provide examples of software applications that engage with complexity in playful ways.

Complexity is beautiful, and beauty - aesthetic pleasure - helps foster engagement. But more importantly, as Smith argues, the pleasures of complexity arise from the discovery of its underlying patterns and structures. Our aim is to display the intrinsic complexity of large collections in ways that enable users to discover those structures. Whereas search-based interfaces assume a specific goal, this process of discovery is more likely to be open-ended, experimental, and driven by curiosity and intrinsic enjoyment. It may, in short, be playful.

References

Bertini, Enrico. 2011. 'Can visualization influence people? I mean can we prove it?' Blog. *Fell in Love with Data*. April 13. http://fellinlovewithdata.com/reflections/can-visualization-influence-people-i-mean-can-we-prove-it.

Braben, David, and Ian Bell. 1984. *Elite*. Acornsoft.

Caillois, Roger. 1961. *Man, Play and Games*. New York: Free Press of Glencoe.

Chen, Jenova. 2007. Flow in games (and everything else). *Communications of the ACM* 50 (April): 31–34. doi:10.1145/1232743.1232769.

Diakopoulos, Nicholas, Funda Kivran-Swaine, and Mor Naaman. 2011. ‘Playable data: characterizing the design space of game-y infographics’. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, 1717–1726. CHI ’11. New York, NY, USA: ACM. doi:10.1145/1978942.1979193.

Ennis Butler, Ben. 2010. *Australian Prints and Printmaking Visualisation*. http://honours.beneb.com/

Ferreira de Oliveira, M. C, and H. Levkowitz. 2003. ‘From visual data exploration to visual data mining: a survey’. *IEEE Transactions on Visualization and Computer Graphics* 9, no. 3 (September): 378- 394. doi:10.1109/TVCG.2003.1207445.

Few, Stephen. 2011. ‘Teradata, David McCandless, and yet another detour for analytics’. Blog. *Visual Business Intelligence*. April 15. http://www.perceptualedge.com/blog/?p=935.

Gell-Man, Murray. 1995. ‘What is complexity’. *Complexity* 1, no. 1: 16–19.

Harris, Jonathan, and Sep Kamvar. 2005. ‘We Feel Fine’. http://www.wefeelfine.org/.

Heylighen, F. 1996. ‘What is complexity?’ *Principia Cybernetica Web*. December. http://pespmc1.vub.ac.be/COMPLEXI.html.

Hinton, Sam, and Whitelaw, Mitchell. 2010. *Commons Explorer*. http://visiblearchive.blogspot.com/2010/03/commonsexplorer.html

Hornbæk, Kasper, and Morten Hertzum. 2011. ‘The notion of overview in information visualization’. *International Journal of Human-Computer Studies* 69, no. 7-8: 509-525. doi:10.1016/j.ijhcs.2011.02.007.

Huizinga, Johan. 1970. *Homo Ludens: A Study of the Play Element in Culture*. London: Temple Smith.

Jones, Matt. 2009. ‘Data as seductive material’. *Slideshare*. http://www.slideshare.net/blackbeltjones/data-as-seductive-material-spring-summit-ume-march09.

Juul, Jesper. 2003. ‘The Game, the Player, the World: Looking for a Heart of Gameness’. In , ed. Marinka Copier and Joost Raessens, 30-45. Utrecht: University of Utrecht.

Keim, D.A. 2002. ‘Information visualization and visual data mining’. *IEEE Transactions on Visualization and Computer Graphics* 8, no. 1 (March): 1-8. doi:10.1109/2945.981847.

Keller, Tanja, and Sigmar-Olaf Tergan. 2005. ‘Visualizing Knowledge and Information: An Introduction’. In *Knowledge and Information Visualization*, ed. Sigmar-Olaf Tergan and Tanja Keller, 3426:1-23. Berlin, Heidelberg: Springer Berlin Heidelberg. http://www.springerlink.com/index/10.1007/11510154_1.

Malaby, Thomas M. 2007. ‘Beyond Play: A New Approach to Games’. *Games and Culture* 2, no. 2 (April 1): 95-113. doi:10.1177/1555412007299434.

Meier, Sid. 1991. *Civilization*. Microprose.

Ramachandran, V. S, and W. Hirstein. 1999. ‘The Science of Art: A Neurological Theory of Aesthetic Experience.’ *Journal of Consciousness Studies*, 6 6.7 (1999) : 15–51. Print.

Shneiderman, Ben. 1996. ‘The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations’. In *Proceedings of the 1996 IEEE Symposium on Visual Languages*, 336. IEEE Computer Society. http://portal.acm.org/citation.cfm?id=834354.

Smith, Peter Frederick. 2003. *The Dynamics of Delight: Architecture and Aesthetics*. London: Routledge.

Taylor, T.L. 2009. ‘The Assemblage of Play’. *Games and Culture*. doi:10.1177/1555412009343576. http://gac.sagepub.com/content/early/2009/08/20/1555412009343576.abstract.

Weaver, Warren. 1948. ‘Science and complexity’. *American scientist* 36, no. 4: 536.

Whitelaw, Mitchell. 2009. *Visible Archive*. http://visiblearchive.blogspot.com/

Wright, Will. 1989. *Sim City*. Maxis.