Computers in Human Behavior 42 (2015) 20–35

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Creative problem solving in digital space using visual analytics

Jacob L. Cybulski *, Susan Keller, Lemai Nguyen, Dilal Saundage

School of Information and Business Analytics, Deakin University, 221 Burwood Highway, Burwood, VIC 3125, Australia

ARTICLE INFO

Article history:
Available online 23 November 2013

Keywords:
Business analytics
Interactivity
Data visualization
Digital creativity
Visual metaphor

ABSTRACT

This article presents a framework for understanding and explaining digital creativity within the growing area of interactive visual analytics. Through the study of extant literature, existing software products, and our own development experience, various aspects of digital creativity are explored in the context of interactive visual analytics and its application to decision-making and problem-solving.

The proposed framework explores and fuses a number of models of individual, social, and domain creativity. It explains the challenges of the analyst navigating through rapidly growing and ubiquitous digital data with an objective to explore it, discover its meanings and associations, as well as solve problems and arrive at effective business decisions. As a creative process, interactive visual analytics differs from other forms of digital creativity, as it utilizes analytic models, relies on the analyst's mental imagery and involves an iterative process of generation and evaluation of ideas in digital media, as well as planning, execution, and refinement of the associated actions. This process is also characterized as collaborative and social by nature as it comprises of analysts from data, problem, and visual domains, who share ideas and actions during analytic activities.

We conclude by suggesting that interactive data visualization may provide opportunities for lay people to creatively engage with data analytics to explore the vast data resources that are freely available and in so doing, gain and communicate insights which may have the potential to impact their private lives and the world at large.

1. Introduction

Interactive visual analytics (IVA) is a visual approach to analytical reasoning, which has set in motion a paradigm shift for handling massive, open, and complex data. IVA combines human perception, cognition, and interaction with analytics to process massive data sets (Arias-Hernández, Dill, Fisher, & Green, 2011; Brodbeck, Mazza, & Lalanne, 2009; van Wijk, 2005). Our society is truly immersed in the digital world. Digital "natives" generate and consume a vast amount of data in a variety of formats at breakneck speed. For example, the amount of data generated annually grew from 150 exabytes in 2005 to 1200 exabytes in 2010 (Helbing & Balietti, 2011) and it is predicted to grow by 40% every year in the near future (Manyika et al., 2011). The huge quantity of readily available data, which is also referred to as "big data", comes in many different forms including text, web data, tweets, sensor data, audio, video, click streams and log files from sensors, smart devices, and social collaboration technologies (Schroeck, Shockley, Smart, Romero-Morales, & Tufano, 2012). Electronic instrumentation and web connectivity continually generate data at a speed that conventional systems can no longer capture, store, and analyze (Schroeck et al., 2012).

The worldwide data explosion raises a key challenge as to how to make sense of big data. This problem requires novel analytical approaches to examining big data, especially because much of the newly generated data is not well understood. Large parts of freely available data are unstructured and often describe human activities that are inherently unpredictable. These data are difficult to analyze with simple graphing and non-interactive techniques. It is also a challenge to analyze relationships between data elements that are not explicit and often not known up front. The very nature of big data provides a strong case for the pursuit of novel analytical approaches capable of creatively combining formal data analysis, rich data visualization, use of familiar metaphors for data representation, and interactive manipulation of these representations, leading to comprehensible reports (Evelson & Yuhanna, 2012).

In order to develop a framework to understand the role of creativity in IVA, in this article we explore several dimensions of creativity, such as those related to innovative products (Amabile, 1983), creative process (Shneiderman, 2000), and domain specific and general creativity (Baer & Kaufman, 2005), the role of intrasubjectivity (e.g. Perry-Smith & Shalley, 2003) and especially the need for intersubjectivity (Drazin, Glynn, & Kazanjian, 1999) in the context of IVA.

* Corresponding author. Tel.: +61 3 9244 6847.
E-mail address: jacob.cybulski@deakin.edu.au (J.L. Cybulski).

0747-5632/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.chb.2013.10.061

This article is organized as follows: Section 2 provides an overview of IVA; Section 3 discusses the framework to understand the role of digital creativity in IVA; Section 4 summarizes the article, its contributions and future directions.

2. Big data and the need for visualization

IVA allows analysts to combine computational power of computers with additional perceptive and cognitive abilities to interactively manipulate and derive insights in support of decision-making from big data (Arias-Hernández et al., 2011; Brodbeck et al., 2009; Van Wijk, 2005). The analysts' engagement with data visualizations plays a significant role in the discovery and communication of insight in IVA. Such engagement involves iterative analysis of datasets from a variety of different viewpoints and interactive manipulation of data attributes to gain new knowledge that is not readily apparent. In other words, IVA aims to "detect the expected and discover the unexpected" from massive, dynamic, ambiguous, and often conflicting data (D. Keim et al., 2008).

IVA has its roots in the fields of scientific and information visualizations. Scientific visualization focuses on the study of large pools of scientific data from sensors, simulations or laboratory tests; while information visualization deals with communication of abstract data and providing the capabilities to transform this data through the use of interactive interfaces (D. Keim et al., 2008). In the past, neither of these visualizations dealt specifically with decision-making, whereas IVA uses decision sciences and analytical sciences to directly assist in decision-making processes based on insights derived from the analysis of big data.

The main premise of IVA is that interactive visualization provides a more effective way to understand and discover new insights from big data, especially in the process of making significant decisions or formulating data-driven action plans (D. Keim et al., 2008). Recent research reports argue that big data has an enormous potential to provide better outcomes to individuals, businesses, and society. Chen, Chiang, and Storey (2012) highlight e-commerce and market intelligence, e-government and politics, science and technology, smart health and wellbeing, and security and public safety as the areas that have the potential to benefit most from big data applications. For example, Google is involved in analyzing queries submitted to its search engine to predict flu outbreaks (Ginsberg et al., 2009) and SAS is involved in predicting unemployment figures from social media sites (SAS, 2011) before officially collected government data is analyzed and reported. Similarly, telecommunication companies study data obtained from location services on mobile devices to make inferences about people and the economy. Medical research organizations undertake examination of health records with a view to identifying useful links between medical treatments and health outcomes (Bollier, 2010).

An example of a simple IVA application is an IBM Many Eyes visualization of movie box office data (see Fig. 1). The objectives of this visualization are shown in Table 1. The visualization classifies movies based on their production budget and Internet Movie Database (IMDB) ratings. In the treemap chart, small rectangles represent individual movies with the size indicating production budget and color specifying the IMDB rating for that particular movie (gray designates poor rating and orange good rating), and the larger rectangular collections of movies denote cinematic genres. One quick insight that can be drawn from this visualization is that not all big budget actions movies are popular with movie goers (e.g. see the selected movie in Fig. 1); and another is that adventure movies tend to have large production budgets and better ratings. The construction of such a chart is a creative act of balancing the chart functionality against its esthetics, while considering a range of constraints imposed on the interactivity and visual options. Playful data explorations by casual users can also constitute the creative pursuit of knowledge and insight.

The general direction of analyzing big data is to sift through data in new ways in search of hidden patterns, relationships, and knowledge that we could not otherwise find (Bollier, 2010). In many instances, visualization of data is preferred over numerical data summaries. Brodbeck et al. (2009, p. 30) define data visualization as a "pipeline that transforms raw data into images that can be interpreted by the human perceptual system". However, as Brodbeck and colleagues suggest, often the main purpose of visualization is not just understanding of data, but rather enabling the analysts to experience data in new ways, and thus appeal to their intuition.

Visualization is also considered a key element in problem-solving. For example, Van Wijk (2005, p. 81) suggests that in a problem space, visualization should support the analyst in extracting relevant information from data, gaining new and surprising insights, as well as making decisions and formulating useful action plans. Van Wijk, however, cautions the potential users of this technology that data visualizations could be misleading (e.g. when the scale of a visual form does not relate to the scale of the represented data effect) and easily misused (either willingly or unwillingly).

Data visualization serves not only as an exploratory tool for data understanding and deriving insights and new knowledge but can also be used as a presentation tool for the purpose of illustration, explanation, and communication of results. The main difference between these two data visualization objectives is that a single presentation graphic will be generated for thousands of potential readers while thousands of exploratory visualizations may be designed to support the data investigation by a single analyst (C. Chen, Härdle, & Unwin, 2008).

One of the main problems with existing tools for visualizations is that most of the data analysis systems use interaction metaphors developed decades ago (Cohen, Dolan, Dunlap, Hellerstein, & Welton, 2009; D. Keim et al., 2008; Schroeck et al., 2012) and the adequacy of their visual capabilities cannot match the demands of big data analysis. For example, tools from software companies such as SAS (e.g. SAS Enterprise Miner) and IBM (e.g. IBM SPSS), which for years have been considered state-of-the-art in statistical analysis of business data, are less focused on supporting visual inspection of data. Only recently have these traditional vendors of analytics software been vigorously pursuing novel ways to enhance their data presentation tool-kits to support the analysis of big data by intuitive decision-makers (e.g. SAS Visual Analytics and IBM Many Eyes).

Visualization techniques for data are no longer confined to science applications. Recent surveys (see e.g. D. A. Keim, Mansmann, Schneidewind, Thomas, & Ziegler, 2008; Russom, 2011; Schroeck et al., 2012) suggest that many practitioners are using advanced data visualization techniques to analyze big data while solving medical, business, and engineering problems. Such techniques are becoming a popular tool for analyzing domain data and communication of insights. We are also witnessing a trend in the growth of the general public's interest in the artistic, novel, and creative presentation of big data. Well-known newspapers, such as *The New York Times* and *The Washington Post*, seem to be on the pursuit of innovative visualizations and graphic presentation of facts of interest to the general community, where the intuitive visual forms demonstrate the significance of otherwise "dry" numbers (Bollier, 2010).

IVA is a complex analytic activity supporting decision processes and problem-solving, in which creativity is required and exercised to analyze a massive, messy, and ubiquitous space of digital data. Data visualization can assist analysts in exploring and navigating through this digital space; mentally processing and interacting