Conducting Qualitative Research with Image Data: A Novel Approach for Exploration and Sensemaking among Sport-Related Images Posted to Social Media

Orland Hoeber, University of Regina
Larena Hoeber, University of Regina
Manali Gaikwad, University of Regina

Method - Qualitative (Other)
20-minute oral presentation (including questions) Saturday, June 9, 2018
Abstract 2018-265 8:55 AM
Room: Acadia B

The use of social media has grown in popularity and dispersion over the last decade due to the development and ongoing refinement of multimedia, networking, and mobile technologies. Platforms such as Twitter, Instagram, and Facebook have changed the way people collect, share, search for, and learn from the information available to them. While early versions of social media were predominantly textual, with a strong focus on making social connections, shared bookmarking, microblogging, sharing and discussing news, and collaborative authoring (McCay-Peet & Quan-Haase, 2017), the inclusion of media such as images and short videos is becoming increasingly popular.

Twitter’s microblogging service serves as a special case for the use of non-textual media, with many people including images within their tweets as a way of working around the 140-character barrier imposed by the platform.

Within the context of sport, images serve as a powerful communication medium, allowing stakeholders to convey complex ideas within the confines of social media. For example, fans may post images at an event to illustrate the significance of their fandom, organizers may post images to raise awareness or promote the positive aspects of their event, and sponsors may post images to promote brand and product visibility, and to make positive associations of their brand with the sport assets. The challenge that sport management researchers face when trying to study the use of images within social media is how to collect, organize, explore, and analyze such data (see Hand, 2017). Previous sport management research using images within social media data (e.g., Clavio & Eagleman, 2011; Guerin-Eagleman & Burch, 2016) has primarily relied on a manual content analysis approach. While this allows for hands-on connection with the data, researchers face criticisms related to completeness and exhaustiveness of their sample sizes (Hand, 2017). The value of addressing these challenges is that doing so “presents new opportunities for drawing upon visual data to enrich our understanding of social media” (Hand, 2017, p. 216).

The purposes of this presentation are to discuss the challenges of conducting qualitative research on image data posted on Twitter, and to present a software tool we have developed to enable such research, called ImgSEE (Image SEarch and Exploration). Given the information-centric nature of this problem, we frame our research in the context of information behaviour (Wilson, 1999), with particular focus on exploratory search (White & Roth, 2009) and sensemaking (Pirolli & Russell, 2011). Following the lead of Tinati, Halford, Carr, and Pope, we integrate “technical capabilities with in-depth qualitative research methods” (2014, p. 6), providing “slightly different approaches” to standard social science analysis methods (Rasmussen Pennington, 2017, p. 232). We leverage information visualization techniques (Ware, 2013) and support for interactive information retrieval (Belkin, 2015) in order to enable the co-development of research questions and the mechanisms for isolating relevant images for answering these research questions.

ImgSEE collects data from Twitter using researcher-defined harvest queries. While Twitter will provide all tweets that make use of these query terms, our interest is only in those tweets that have embedded images. The images are downloaded in a separate process, and analyzed using image processing and deep learning algorithms. The extracted information includes the average colour values calculated from the low-level image pixels, textual contents within the image obtained using optical character recognition, and a classification of the image against a library of 1000 generic tags that were used to train a neural network. In addition, the metadata (e.g., text) associated with the tweet that contained the image is maintained in association with the image.

The search interface provided by ImgSEE leverages this extracted information to enable information seeking within the image and associated data. Researchers may issue queries within the harvested data, and can control whether
these queries are matched against the textual contents of the tweet, text within the images, and the machine-derived tags. While traditional image search interfaces organize images within a grid-based structure, the mechanism for ranking the images does not produce a consistent and learnable structure. Instead, ImgSEE uses the ambient colour within the image as a mechanism for organizing the images based on their colour similarity. These are organized within a disc, with a perceptually-ordered colour scale presented at the outside edge of the disc. Within this colour-space, images with similar colours are grouped together, enabling the researcher to observe the overall colour makeup of the image set, and supporting interactive pan and zoom within the image space. A timeline of when the images were posted is provided, and the top hashtags associated with the images are presented along with micro-visualizations of the colour-spaces associated with the corresponding images. These interface features work as coordinated views of the image data, such that filtering or focusing in one view results in corresponding changes in the other views.

The ability to dynamically query the image dataset based on various different textual attributes, the organization of the images within a zoomable image-space, and the tools that support interactive filtering based on temporal and topical (hashtag) attributes support exploratory search within this data. Such a high degree of interactivity has been shown in other work to support the interactive production of purposive samples for qualitative research, and the co-development of research questions during the exploratory search process (Hoeber, Hoeber, Snelgrove, & Wood, 2017).

In order to further support the qualitative analysis of the image data, a workspace is provided that enables the saving of potentially relevant images, and the dynamic organization of these images within researcher-defined categories. Being able to drag-and-drop images into multiple different categories enables interactive sensemaking and preliminary data coding during the exploration process. These saved images can be downloaded to the researcher’s computer through an export function, with the category names becoming the directories in which the images are saved. The goal here is to provide a smooth transition from exploratory search and sensemaking to detailed content analysis and contextualization of the image collections (Hand, 2017).

Although social media provides a convenient platform for collecting images, conducting qualitative research on such data is challenging due to the potentially overwhelming amount of image data that can be collected in a short timeframe. Our work provides sport management researchers with tools and techniques to enable this type of research, the value of which will be illustrated through an example from images posted to Twitter during the 2017 Le Tour de France. We will conclude the presentation with a discussion on how these approaches may be extended to other image-rich settings, such as print media, marketing and sponsorship materials, and historical documents.