IUI-TextVis 2015 Fourth Workshop on Interactive Visual Text Analytics

Jaegul Choo
Georgia Institute of Technology
jaegul.choo@cc.gatech.edu

Christopher Collins
University of Ontario Institute of Technology
christopher.collins@uoit.ca

Wenwen Dou
University of North Carolina at Charlotte
wdou1@uncc.edu

Alex Endert
Georgia Institute of Technology
endert@gatech.edu

ABSTRACT

Analyzing text documents has been a key research topic in many areas. Countless approaches has been proposed to tackle this problem, and they are largely categorized into fully automated approaches (via statistical techniques) or human-involved exploratory ones (via interactive visualization). The primary purpose of this workshop is to bring together researchers from both sides and provide them with opportunities to discuss ways to harmonize the power of these two complementary approaches. The combination will allow us to push the boundary of text analytics. The detailed workshop schedule, proceedings, and agenda will be available at `http://www.textvis.org`.

Author Keywords

text analysis; visual text analytics; heterogeneous text; exploratory text mining; interactive visual text mining

ACM Classification Keywords

H.1.2 Information Systems: Models and Principles—Human factors

MAIN OBJECTIVES

In the big data era we are entering, analyzing free-text documents is one of the biggest challenges. A myriad of automated approaches from machine learning, data mining, information retrieval, and natural language processing have been proposed so far, while the research fields such as human-computer interaction, information visualization, and visual analytics have also been developing their own interactive methods and systems.

IUI 2015, March 29–April 1, 2015, Atlanta, GA, USA.
ACM 978-1-4503-3306-1/15/03.
http://dx.doi.org/10.1145/2678025.2716270

The primary goal of the workshop[1] is to bridge the gap by bringing together researchers from both sides, which allows us to push the boundary of text analytics. The workshop should provide an opportunity to discuss and explore ways to harmonize the power of automated techniques and the exploratory nature of interactive visualization. Built on the theme of the previous workshops successfully held in 2011-2013 in conjunction with IEEE VIS, this workshop at ACM IUI'15 will foster ideas from a multi-disciplinary community of researchers to address the topics outlined below.

Specifically, the workshop will investigate the following challenges and discuss promising directions in text analytics.

Heterogeneous and ubiquitous text analytics. In many cases, text data are associated with other data types, such as authors, years, and locations. The workshop will collect various use cases about heterogeneous data and ubiquitous text analytics. From the use cases, one will be able to better understand the requirements of heterogeneous textual data analysis from a task-driven perspective. Although there is some previous work on visual analysis of heterogeneous textual data, there is not a clear understanding of the typical tasks that people would like to achieve in analyzing heterogeneous textual data.

Adaptive and intelligent analytics. The IUI conference offers a unique opportunity to consider how advanced machine learning and artifical intelligence techniques can be leveraged to improve visual text analytics. For example, how can an intelligent analytics interface suggest appropriate starting points for an investigation of a large document collection? Suggested views could be based on models of user knowledge and automated analysis of the dataset and contextual cues. Moving into the realm of IUI for visual text analytics we envision systems in which the interface is responsive to data and user characteristics.

Task-driven text analytics. How will different tasks influence the design and development of both text analytics and visualization technologies? The workshop will leverage the power of the crowd at the workshop to examine a number of

[1] `http://www.textvis.org`

use cases and draft a taxonomy that characterizes the design dimensions of the space and can also be used to guide the future design and development. Second, based on the use cases, the workshop will examine how to best leverage state-of-the-art text analytics and traditional data mining techniques in conjunction with novel interactive visual analytics to address the challenges manifested by the collected use cases.

Interactive visual text mining techniques. Most of the automated text mining and natural language processing algorithms cannot fully understand the semantics of textual data. The workshop will address how to improve them via human-in-the-loop interactive visualization techniques.

Real-time streaming text visual analytics. Nowadays, a lot of textual data from social media such as twitter has real-time aspects. The workshop will present the studies addressing these aspects through real-time text visual analytics systems and associated techniques.

Other than the above-mentioned topics, the workshop will broadly cover other related directions that can advance the state-of-the-art text analytics.

TOPICS OF INTEREST

The topics of interest for the workshop include (but are not limited to) the following:

- User-adaptive approaches to visual text analytics
- Coupling of machine learning- or artificial intelligence-based systems with visualization for text analytics
- Visualization to improve automated algorithms and models for text data
- Coordinated visualizations of textual and non-textual data
- Visual metaphors for textual data or ubiquitous text analytics
- Perception and cognition in visual text analytics
- Systems, languages, and architectures for visual textual analytics
- Collaborative analysis of textual and linked data
- Real-time visualization of streaming textual data
- Visual opinion analysis from textual data
- Visual event identification and prediction from textual data
- Uncertainty in interactive visual textual analytics
- Industry-specific applications of visual analysis of textual data, e.g., retail, healthcare, government, etc.
- Studies and evaluation of textual data visualization techniques, systems, metrics, and bench-marks
- Standardized datasets and tasks for visual text analytics

ORGANIZERS

The organizers and their qualifications are as follows:

Jaegul Choo

Dr. Jaegul Choo is a research scientist in the School of Computational Science and Engineering at Georgia Institute of Technology. He received a Ph.D in computational science ang engineering from Georgia Institute of Technology in 2013. He is broadly interested in integrating data mining techniques with visual analytics. He primarily focuses on dimension reduction and clustering techniques, which play essential roles in visualizing and interacting with large-scale high-dimensional data. Recently, he has mainly worked on text visual analytics research including an interactive topic modeling visual analytics system based on nonnegative matrix factorization published in IEEE Transactions on Visualization and Computer Graphics in 2013. He has published in SDM, WSDM, WWW, DMKD, VAST, IEEE TVCG, IEEE CG&A, etc. and has served as a program committee in various conferences and workshops including VDA, SDM, AISTATS.

Christopher Collins

Dr. Christopher Collins, Canada Research Chair in Linguistic Information Visualization, is an Assistant Professor of Computer Science at the University of Ontario Institute of Technology (UOIT). His research focus is interdisciplinary, combining information visualization and human-computer interaction with natural language processing to address the challenges of information management and the problems of information overload. His visual text analytics research has a focus on designing analytic interfaces for touch and gesture modalities. His publications have appeared in IEEE TVCG, IEEE VIS, EuroVis, ACM ITS, and other venues. Collins received his PhD in Computer Science from the University of Toronto in 2010. He collaborates with several universities, IBMs TJ Watson Research Laboratory, and a number of smaller software firms. He is a member of the IEEE Visualization and Graphics Technical Committee Executive.

Wenwen Dou

Dr. Wenwen Dou is a research assistant professor in the department of computer scientist at UNC Charlotte. Her research interests include Visual Analytics, Text Mining, and Human Computer Interaction. In particular, her research lies at the intersection of interactive visualization and text mining, with the specific area referred to as Visual Text Analytics. She has conducted extensive research on interactive visual analysis of large text corpora, in particular with focus on temporal analysis and event extraction. Dou has worked with various analytics domains in reducing information overload and providing interactive visual means to analyzing unstructured information. She has experience in turning cutting-edge research into technologies that have broad societal impacts, partially demonstrated by supports from both academic and industry partners, including PNNL, US Army Research Office, US Special Operations Command, NSF, US Army Engineering Research and Development Center, and Lowes company Inc.

Alex Endert

Dr. Alex Endert is an assistant professor in the School of Interactive Computing at Georgia Tech. His interests are

primarily on visual data exploration, human-computer interaction, and information visualization. His research centers around exploring novel user interaction techniques for visual analytic systems to attempt to couple cognitive and computational processes for sensemaking. He is an active member of, and contributor to, premier venues for human-computer interaction and information visualization (ACM CHI, IEEE VIS, IEEE TVCG, IEEE CG&A). He received his Ph.D. in Computer Science at Virginia Tech in 2012, where he worked with Dr. Chris North. In 2013, his work on Semantic Interaction was awarded the IEEE VGTC VPG Pioneers Group Doctoral Dissertation Award, and the Virginia Tech Computer Science Best Dissertation Award.

PROGRAM COMMITTEE

The program committee of the workshop will include (not all of them are confirmed at the time of this proposal submission):

- Jason Chuang, University of Washington
- Nick Cramer, Pacific Northwest National Laboratory
- Weiwei Cui, Microsoft Research Asia
- Eser Kandogan, IBM Almaden Research Center
- Zhicheng Liu, Adobe Research
- Naren Ramakrishnan, Virginia Tech
- Christian Rohrdantz University of Konstanz
- John Stasko, Georgia Insitute of Technology

WORKSHOP PROGRAM

The workshop will be held for a full day on 2015/3/29. It will mainly have four sessions: two in the morning and the other two in the afternoon. We plan to organize them as a balanced mixture of one or two keynote talks, oral presentations of accepted papers, poster presentations, and panel discussions. Full details can be found at `http://www.textvis.org`.

ACKNOWLEDGMENTS

The organizers would like to thank all the authors for contributing to the workshop and all the members of the program committee for establishing the scientific quality of the workshop.