DaVInt - Data Visualization and Interaction Lab

Isabel H. Manssour and Milene S. Silveira
Pontifical Catholic University of Rio Grande do Sul, PUCRS.
School of Technology. Graduate Program in Computer Science.
Av. Ipiranga, 6681 - Prédio 32. 90619-900, Porto Alegre, RS, Brazil.
{isabel.manssour,milene.silveira}@pucrs.br
https://www.inf.pucrs.br/davint/

***Abstract*—A multidisciplinary research project between Communication and Computer Science researchers gave rise to the Data Visualization and Interaction Lab, known as DaVInt lab. Officially created in 2017, DaVInt comprises students and researchers from Communication, and Computer Science and related areas. DaVInt's main research topics are related to visual analytics of social media, narrative visualization, and data science. However, computer vision projects and other individual initiatives are also developed. Its main contributions are divided into three main areas: scientific, through our publications; technological, making the results of our researches available to the community; and student education, through which we disseminate the area beyond academia.**

I. Introduction

The DaVInt lab (Data Visualization and Interaction) emerged from a multidisciplinary research project between the School of Communications, Arts, and Design (FAMECOS) and the School of Technology from PUCRS. This project started in 2011 with a focus on social media data visualization. However, just in 2017, with the increase of the team and with the equipment acquired over the years, the lab was officially created. It has a multidisciplinary team composed of students and researchers from communication, computer science and related areas, administration, and engineering.

Although we continue researching the initial topic of social media data visualization, novel research topics were added through these years. Recently, we have dedicated efforts to verify how visualization can help in data preprocessing activities [1]–[4], a task that demands a lot of time for data analysts [5]. A result of this research can be seen in a paper we recently published with a conceptual process that includes preprocessing profiling as a new phase in the Visual Analytics (VA) process [4].

We also work with visualization applied research, with education as one of the areas that have been explored [6]. Other topics of interest are the visualization of geospatial data [7] and the narrative visualization (also known as visual storytelling) [8]. In the first one, we are collaborating with researchers in the field of high-performance processing [9], [10]. In the second one, the research aims to help the design of narrative visualizations, presenting models to represent this kind of visualization [11] and, also, to support their construction, focusing on the perspective of end-user development [12], [13].

In addition to research in data visualization and VA, we work with computer vision. We started by developing research projects aimed at assisting the mobility of visually impaired persons [14]–[17]. However, this year we started researching the use of computer vision for the security area. More specifically, we are working on identifying firearm threats in videos.

We also focus on training people in this area, disseminating the importance of data visualization and VA. Considering that many of our Master's students are from industry, they help us in this dissemination beyond academia and bring real problems that benefit from the visualization area. In this context, we recently had an accepted paper that uses VA for churn reduction [18]. Besides, we have a master's student who is a statistician and works at a fintech, developing a work that uses VA for monitoring credit scoring models to assist in managing models and decision making.

The remainder of the paper is organized as follows. Section II presents our main research projects. The contributions of our research are described in Section III. In the last section, we outline our conclusions, as well as future research possibilities.

II. Main Projects

In the following sections, we briefly describe the research projects under development in the DaVInt lab. Some visualizations resulting from these projects can be seen in Figure 1.

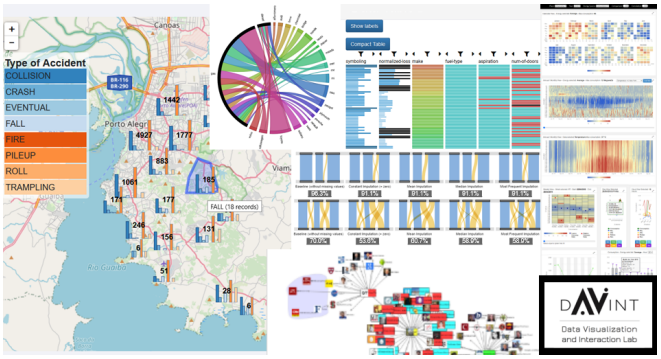

Fig. 1. Images resulting from some projects developed at DaVInt.

A. Visual Analysis of Social Media

This project is the oldest and also the one that gave rise to the laboratory. But the research that is carried out by

a multidisciplinary group has been changing and updating itself over time. In the beginning, the focus was on discourse analysis, which involved collecting Tweets from different subjects and was more focused on data-driven journalism [19], [20].

Then, taking advantage of the knowledge acquired and the tools developed, we worked with case studies related to politics [21], [22]. And, in recent years, we've explored the visual analysis of the effects of audiovisual narratives on streaming data from social networking sites. This research resulted in PeakVis, an interactive tool that syncs a video recording with a responsive line graph representing the total number of tweets from each moment, the top messages, a dynamic word cloud, and a semantic graph showing word correlation [23]. The motivation is that nowadays, television broadcasts are expected to be accompanied by a stream of social media postings by the audience, mediating the asymmetrical relationship between the television network and spectators through commentary, humor, and criticism.

B. Narrative Visualization

Narrative visualizations blend narrative elements and data visualization to both improve data comprehension and the users' involvement in the interaction with such data. Our research on this topic intends to support designers and end-users in designing this kind of visualization.

We started trying to understand the narrative design space - the design strategies used in narrative visualizations - to enable the identification of ways to design clear and concise narrative visualizations [8]. Then, in the context of social media, we proposed a model for the representation of narratives through visualizations, intending to support the construction of interactive narratives, contemplating not only the construction of the interface itself but also the mapping of the data to its visual representation [11]. After that, we proposed a customizable interface design model, with the main purpose of assist designers and data analysts as co-designers to reflect on different aspects that influence the construction of customizable interactive narrative visualizations [12]. Now, we are investigating the relationship between the elements that compose a narrative and different data types and deepen the studies about end-user development in this field.

C. Computer Vision

There are more than 35.7 million people with visual impairment in Brazil, and the resources used to support their mobility in urban environments, e.g., white cane, offer little information. Although there are studies related to navigation support for those individuals, there have still been gaps to be researched. For a few years, we explored these gaps in developing a model [16] that, based on computer vision techniques, can aid in the mobility of these people. This model offers an integrated solution for localization and identification of tactile paving surfaces [14], detection of aerial and ground obstacles, and localization of crosswalks [15]. We also used deep neural networks for addressing: the development of an approach to classify different scenarios of indoor environments with or without doors and stairs [24]; the detection of pedestrian traffic lights together with their current state for helping visually-impaired people to cross the streets with the aid of their mobile devices [17].

We also recently used deep neural networks for Braille letters detection to assist in teaching programming for blind students using a scaled-down physical environment [25]. Furthermore, currently, we are interested in helping in the security area, exploring one of the most commonly used security measures: security cameras. Aiming to identify when a threatening event occurs, we proposed a new dataset for identifying firearm threats. We are now developing a simplified version of neural network architectures to efficiently identify firearm threat events in videos captured by a cellphone or security camera.

D. Data Science

In a study of more than 300 respondents working in fields related to visual analytics, presented by Forbes Insights 2017[1], only 42% of respondents were confident in the quality of their data. Furthermore, data visualization has become one of the foundations in KDD (knowledge discovery in databases) and data science since this process derives significant insights into the data under analysis [5], [26], [27].

Concerned about assisting in data quality analysis, we started by conducting semi-structured interviews with thirteen data analysts. The responses of the interviewees led us to a list of ten insights for new visualization solutions [1]. After that, we developed a conceptual VA process that includes preprocessing profiling as a new phase, and we indicate a list of research opportunities in the scope of preprocessing combined with visualization and VA [4]. This work was developed with the collaboration of Fernando Paulovich[2] from Dalhousie University. We continue to explore this line of research at DaVInt with some master's students.

E. Other Initiatives

We also work at DaVInt lab with applied research or some alone initiatives. Some are the result of the need to meet some isolated funding, and others arise according to the specific interests of some students. One example is the visual analytics system for energy data to support decision-making in smart cities and buildings [28]. It allows the simultaneous analysis of meteorological data and consumption or generation energy data sources, such as air conditioners and solar energy, respectively. Besides statistical functions for data understanding, it also has prediction algorithms to foresee energy data based on past energy data patterns. Another example is the GeoMultiVis [7], a set of interactive visualizations that combines multivariate data and geospatial localization. It allows to apply different filters interactively and analyze

[1]https://i.forbesimg.com/forbesinsights/d&b_enterprise_analytics/Analytics_Accelerates_Into_Mainstream.pdf

[2]https://www.dal.ca/faculty/computerscience/faculty-staff/fernando-paulovich.html

several attributes, helping to find patterns and insights that could support decision-makers.

The use of visualization and VA to help the education area is another area of research explored in the laboratory. We started exploring how interaction and visualization techniques integrated with data mining algorithms can assist teachers in predicting students' performance in learning environments. Then, we developed a visual analysis approach to investigate and predict pass/fail rates in distance learning courses [6], [29]. We are now researching the development of visualization-based solutions to support self-regulatory processes for online learning. This research led us to the development of a digital storytelling-based solution for improving learning motivation and engagement.

The DaVInt lab also includes research topics more closely related to the area of Human-Computer Interaction, but not related to the specific focus of this Conference.

III. Contributions

We believe that our contributions are divided into three main areas: scientific, technological, and student education. Scientific contributions can be verified through our publications, such as those mentioned in the previous section.

Concerning technological contribution, recently, we began to dedicate ourselves to making the results of our researches available to the community. We organize the software on GitHub for use in the area of data quality analysis[3][4], and social media analysis[5][6]. Prototypes resulting from master's dissertations are also available[7][8].

The student's education occurs at different levels, from graduation, including students often participating in their first scientific initiation program, up to Ph.D. students in the final years of the graduate course. The different levels of education combined with the group multidisciplinarity enrich group meetings and discussions and, consequently, the student's and researcher's development. After graduation, students take positions in both industry and academia.

IV. Final Remarks

Currently - regarding specifically the WVIS topics - we intend to continue researching the use of visual analytics to understand the data under analysis better. Our focus is to gain insights from a collected dataset and to help in the detection of data quality issues that are important for being adjusted in the first phases of data analysis, considering the Data Science context. The use of visual analytics toward other research topics, such as the projects presented in Section II, is also of interest and continues to be researched.

Beyond our established cooperations, we are also open to novel collaborations with research groups from Brazil and abroad, which are researching topics related to the ones presented here or that are interested in our projects.

[3]https://github.com/DAVINTLAB/pandas-profiling

[4]https://github.com/DAVINTLAB/preprocessing-profiling

[5]https://github.com/DAVINTLAB/TweetUtils

[6]https://github.com/DAVINTLAB/Peakvis

[7]https://github.com/DAVINTLAB/GeoMultiVis

[8]https://github.com/DAVINTLAB/VA-EnergyData

Acknowledgment

The authors would like to thank PUCRS for funding the laboratory. Isabel Harb Manssour also would like to thank the financial support of the CNPq Scholarship - Brazil (308456/2020-3).

References

[1] A. M. P. Milani, F. V. Paulovich, and I. H. Manssour, "Visualization in the preprocessing phase: Getting insights from enterprise professionals," *Information Visualization*, vol. 19, no. 4, pp. 273–287, 2020. [Online]. Available: https://doi.org/10.1177/1473871619896101

[2] L. Ciocari, A. Milani, and I. Manssour, "Uso de visualização de dados para auxiliar na análise e pré-processamento de dados categóricos," in *Anais Estendidos do XXXIII Conference on Graphics, Patterns and Images*. Porto Alegre, RS, Brasil: SBC, 2020, pp. 195–198. [Online]. Available: https://sol.sbc.org.br/index.php/sibgrapi_estendido/article/view/13010

[3] A. M. Milani, F. Paulovich, and I. Manssour, "Preprocessing profiling model for visual analytics," in *Anais Estendidos do XXXIII Conference on Graphics, Patterns and Images*. Porto Alegre, RS, Brasil: SBC, 2020, pp. 105–111. [Online]. Available: https://sol.sbc.org.br/index.php/sibgrapi_estendido/article/view/12991

[4] A. M. P. Milani, L. A. Loges, F. V. Paulovich, and I. H. Manssour, "Prava: Preprocessing profiling approach for visual analytics," *Information Visualization*, vol. 20, no. 2-3, pp. 101–122, 2021. [Online]. Available: https://doi.org/10.1177/14738716211021591

[5] R. Jugulum, *Competing with high quality data: Concepts, tools, and techniques for building a successful approach to data quality*. John Wiley & Sons, 2014.

[6] A. Weiand, I. H. Manssour, and M. S. Silveira, "Visual analysis for monitoring students in distance coursess," *International Journal of Distance Education Technologies (IJDET)*, vol. 17, no. 2, pp. 18–44, 2019. [Online]. Available: http://doi.org/10.4018/IJDET.201904010

[7] W. Signoretti, I. H. Manssour, and M. S. Silveira, "Geomultivis: helping decision-making through interactive visualizations from geospatial multivariate data," *Proceedings of AMCIS 2019, 2019, Brasil.*, 2019.

[8] E. Ghidini, C. Q. Santos, I. Manssour, and M. S. Silveira, "Analyzing design strategies for narrative visualization," in *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems*, ser. IHC 2017. New York, NY, USA: Association for Computing Machinery, 2017. [Online]. Available: https://doi.org/10.1145/3160504.3160538

[9] C. Ledur, D. Griebler, I. Manssour, and L. G. Fernandes, "A high-level dsl for geospatial visualizations with multi-core parallelism support," in *2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC)*, vol. 1, 2017, pp. 298–304.

[10] E. Ewald, A. Vogel, C. Rista, D. Griebler, I. Manssour, and L. Gustavo, "Parallel and distributed processing support for a geospatial data visualization dsl," in *2018 Symposium on High Performance Computing Systems (WSCAD)*, 2018, pp. 221–228.

[11] E. Ghidini, "Representação de narrativas interativas por meio de visualizações com dados extraídos de redes sociais," Master's thesis, Pontifical Catholic Univeristy of Rio Grande do Sul (PUCRS), 2017.

[12] C. Q. Santos, "Civis: modelo de design de interface customizável para apoiar a construção de visualizações narrativas interativas de dados extraídos de mídias sociais," Ph.D. dissertation, Pontifical Catholic Univeristy of Rio Grande do Sul (PUCRS), 2018.

[13] C. Q. Santos and M. S. Silveira, "Qualitative studies on human-data interaction and data analysis collaboration in latin america," in *Proceedings of the IX Latin American Conference on Human Computer Interaction*, 2019, pp. 1–4.

[14] M. C. Ghilardi, R. C. Macedo, and I. H. Manssour, "A new approach for automatic detection of tactile paving surfaces in sidewalks," *Procedia Computer Science*, vol. 80, pp. 662–672, 2016, iCCS-Int. Conf. on Computational Science, 6-8 June 2016, San Diego, California, USA.

[15] M. C. Ghilardi, J. C. S. J. Junior, and I. H. Manssour, "Crosswalk localization from low resolution satellite images to assist visually impaired people," *IEEE Computer Graphics and Applications*, vol. PP, no. 99, 2016.

[16] M. C. Ghilardi and I. H. Manssour, "Support model for navigation on sidewalks for visually impaired persons," in *Conference on Graphics, Patterns and Images (SIBGRAPI)*. Porto Alegre, RS, Brasil: SBC, 2016. [Online]. Available: http://urlib.net/rep/8JMKD3MGPAW/3M9HE2L

[17] M. C. Ghilardi, G. Simões, J. Wehrmann, I. H. Manssour, and R. C. Barros, "Real-time detection of pedestrian traffic lights for visually-impaired people," in *2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–8.

[18] P. A. Karolczak and I. H. Manssour, "Using visual analytics to reduce churn," in *To be presented at International Conference on Computational Science and Its Applications (ICCSA)*, 2021.

[19] C. Q. Santos, R. Tietzmann, M. Trasel, S. Moraes, I. H. Manssour, and M. S. Silveira, "Can visualization techniques help journalists to deepen analysis of twitter data? exploring the "germany 7 x 1 brazil" case," in *Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS).*, 2016, pp. 1939–1948.

[20] C. Q. Santos, H. S. da Cunha, C. Teixeira, D. R. de Souza, R. Tietzmann, I. Manssour, M. S. Silveira, M. Trasel, D. D. A. Ruiz, and R. C. Barros, "Media professionals? opinions about interactive visualizations of political polarization during brazilian presidential campaigns on twitter," in *Proceedings of the 50th Hawaii International Conference on System Sciences (HICSS).*, 2017.

[21] C. R. G. Teixeira, G. Kurtz, L. P. Leuck, R. Tietzmann, D. R. de Souza, J. A. M. F. Lerina, I. H. Manssour, and M. S. Silveira, "Humor, support and criticism: A taxonomy for discourse analysis about political crisis on twitter," in *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age*, ser. dg.o 2018. New York, NY, USA: Association for Computing Machinery, 2018, p. 1–6. [Online]. Available: https://doi.org/10.1145/3209281.3209330

[22] C. R. G. Teixeira, G. Kurtz, L. P. Leuck, P. Sanvido, J. Scherer, R. Tietzmann, I. Manssour, and M. Silveira, "Polls, plans and tweets: An analysis of the candidates' discourses during the 2018 brazilian presidential election," in *Proceedings of the 20th Annual International Conference on Digital Government Research*, ser. dg.o 2019. New York, NY, USA: Association for Computing Machinery, 2019, p. 439–444. [Online]. Available: https://doi.org/10.1145/3325112.3329882

[23] P. H. M. Sanvido, G. B. Kurtz, C. R. G. Teixeira, P. P. Wagner, L. Leuck, M. Silveira, I. H. Manssour, and R. Tietzmann, "Peakvis: a visual analysis tool for social network data and video broadcast," in *Presented at IEEE Computer Society Computers, Software, and Applications Conference (COMPSAC)*, 2021.

[24] R. S. N. JR, M. de Richter Herrmann, and I. Manssour, "Recognition of doors and stairs using convolutional neural networks to assist visually impaired persons," in *Anais do XIV Workshop de Visão Computacional*, 2018, pp. 166–171.

[25] D. Gonçalves, G. Santos, M. Campos, A. Amory, and I. Manssour, "Braille character detection using deep neural networks for an educational robot for visually impaired people," in *Anais do XVI Workshop de Visão Computacional*. Porto Alegre, RS, Brasil: SBC, 2020, pp. 123–128. [Online]. Available: https://sol.sbc.org.br/index.php/wvc/article/view/13492

[26] S. Liu, D. Maljovec, B. Wang, P.-T. Bremer, and V. Pascucci, "Visualizing high-dimensional data: Advances in the past decade," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 3, pp. 1249–1268, 2017.

[27] A. Crisan, B. Fiore-Gartland, and M. Tory, "Passing the data baton : A retrospective analysis on data science work and workers," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1860–1870, 2021.

[28] Ânderson Pinto Alves, A. M. P. Milani, and I. H. Manssour, "Visual analytics system for energy data in smart cities and buildings," in *2020 IEEE International Smart Cities Conference (ISC2)*, 2020, pp. 1–8.

[29] A. Weiand and I. H. Manssour, "Towards visual analysis techniques for monitoring students of distance education courses," in *Proceedings of the 6th Workshop on Visual Analytics, Information Visualization and Scientific Visualization*, 2015.