

Presentation of the paper “An architectural proposal to explore the data of a private community through visual analytic”

Jorge Durán-Escudero
University of Salamanca
jorge.d@usal.es

Abstract

This is the presentation of the paper entitled “An architectural proposal to explore the data of a private community through visual analytic” in the Learning Analytics Track of the TEEM 2017 International Conference held in Cádiz (Spain) in October 18-20, 2017.

In this document, a proposal is made to study the data that will be generated in the private and anonymous community of the WYRED project, in order to extract knowledge about how their users interact, both between them, and with the platform. To do this, it is started with the creation of a system that will generate a set of test data, as close as possible to the original. With this information and considering the impact of privacy when dealing with the data of the project, a flexible and complete architecture has been proposed for the development of interactive visualizations that will allow to visualize the previously generated data. Finally, a use case is presented where the suitability of the visual analytic is demonstrated to perform analysis of the data of the project and to extract knowledge, in a simple way.

The presented paper may be cited as:

Durán-Escudero, J., García-Peñalvo, F. J., & Therón-Sánchez, R. (2017). An architectural proposal to explore the data of a private community through visual analytic. In J. M. Doderó, M. S. Ibarra Saiz, & I. Ruiz Rube (Eds.), *Proceedings of the Fifth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'17) (Cádiz, Spain, October 18-20, 2017)* (Article 48). New York, NY, USA: ACM. doi:10.1145/3144826.3145398

Link to the presentation

<https://es.slideshare.net/grialusal/2017-wyred-teem>

Keywords

WYRED; Software Architecture; Visual Learning Analytics

References

- [1] J. Schultz. 2016. Title. Micro Focus Blog. Stories and updates from our team, partners and supporters. <https://blog.microfocus.com/how-much-data-is-created-on-the-internet-each-day/>
- [2] F. J. García-Peñalvo. 2016. Technological Ecosystems. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje* 11, 1, 31-32. DOI:10.1109/RITA.2016.2518458.

- [3] F. J. García-Peñalvo. 2016. The WYRED Project: A Technological Platform for a Generative Research and Dialogue about Youth Perspectives and Interests in Digital Society. *Journal of Information Technology Research* 9, 4, vi-x.
- [4] F. J. García-Peñalvo and N. A. Kearney. 2016. Networked youth research for empowerment in digital society. The WYRED project. In *Proceedings of the Fourth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'16) (Salamanca, Spain, November 2-4, 2016)*, F.J. García-Peñalvo Ed. ACM, New York, NY, USA, 3-9. DOI:10.1145/3012430.3012489.
- [5] A. García-Holgado and F. J. García-Peñalvo. 2013. The evolution of the technological ecosystems: An architectural proposal to enhancing learning processes. In *Proceedings of the First International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'13) (Salamanca, Spain, November 14-15, 2013)*, F.J. García-Peñalvo Ed. ACM, New York, NY, USA, 565-571. DOI:10.1145/2536536.2536623.
- [6] A. García-Holgado and F. J. García-Peñalvo. 2014. Architectural pattern for the definition of eLearning ecosystems based on Open Source developments. In *Proceedings of 2014 International Symposium on Computers in Education (SIIE), Logrono, La Rioja, Spain, 12-14 Nov. 2014*, J.L. Sierra-Rodríguez, J.M. Doderobeardo and D. Burgos Eds. Institute of Electrical and Electronics Engineers, USA, 93-98. DOI:10.1109/SIIE.2014.7017711.
- [7] A. García-Holgado and F. J. García-Peñalvo. 2016. Architectural pattern to improve the definition and implementation of eLearning ecosystems. *Science of Computer Programming* 129, 20-34. DOI:10.1016/j.scico.2016.03.010.
- [8] F. J. García-Peñalvo and J. Durán-Escudero. 2017. Interaction design principles in WYRED platform. In *Learning and Collaboration Technologies. Technology in Education. 4th International Conference, LCT 2017. Held as Part of HCI International 2017, Vancouver, BC, Canada, July 9–14, 2017. Proceedings, Part II*, P. Zaphiris and A. Ioannou Eds. Springer International Publishing, Switzerland, 371-381. DOI:10.1007/978-3-319-58515-4_29.
- [9] D. A. Keim, J. Kohlhammer, G. Ellis, and F. Mansmann (Ed.). 2010. *Mastering the Information Age Solving Problems with Visual Analytics*. Eurographics Association, Goslar, Germany.
- [10] R. Richards. 2015. *Software architecture patterns*. O'Reilly Media, Sebastopol, CA.
- [11] K. Matković, W. Freiler, D. Gračanin, and H. Hauser. 2008. ComVis: A coordinated multiple views system for prototyping new visualization technology. In *12th International Conference Information Visualisation, IV08*, London, 215-220. DOI:10.1109/IV.2008.87.
- [12] Z. A. Pardos and K. Kao. 2015. MoocRP: An open-source analytics platform. In *2nd ACM Conference on Learning at Scale, L@S 2015* Association for Computing Machinery, Inc, 103-110. DOI:10.1145/2724660.2724683.
- [13] F. Huang, C. X. Wang, and J. Long. 2011. Design and Implementation of Single Sign on System with Cluster CAS for Public Service Platform of Science and Technology Evaluation. In *2011IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*, 732-737. DOI:10.1109/TrustCom.2011.95.
- [14] J. Pokorny. 2013. NoSQL databases: a step to database scalability in web environment. *International Journal of Web Information Systems* 9, 1, 69-82. DOI:<https://doi.org/10.1108/17440081311316398>.

- [15] R. Cattell. 2011. Scalable SQL and NoSQL data stores. *ACM SIGMOD Record* 39, 4, 12-27. DOI:10.1145/1978915.1978919.
- [16] L. Sweeney. 2002. k-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10, 5, 557-570.
- [17] D. M. Blei, A. Y. Ng, and M. I. Jordan. 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research* 3, 993-1022.
- [18] J. Boyd-Graber and D. M. Blei. 2009. Multilingual topic models for unaligned text. In *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence (UAI '09) Montreal, Quebec, Canada — June 18 - 21, 2009* AUAI Press, Arlington, Virginia, United States, 75-82.
- [19] J. Jagarlamudi and H. Daumé. 2010. Extracting Multilingual Topics from Unaligned Comparable Corpora. In *Advances in Information Retrieval. ECIR 2010*, C. Gurrin, Y. He, G. Kazai, U. Kruschwitz, S. Little, T. Roelleke, S. Rüger and K. Van Rijsbergen Eds. Springer, Berlin, Heidelberg, 444-456. DOI:https://doi.org/10.1007/978-3-642-12275-0_39.
- [20] K. Kucher and A. Kerren. 2015. Text visualization techniques: Taxonomy, visual survey, and community insights. In *2015 IEEE Pacific Visualization Symposium (PacificVis)*, 117-121. DOI:10.1109/PACIFICVIS.2015.7156366.
- [21] S. Havre, E. Hetzler, P. Whitney, and L. Nowell. 2002. ThemeRiver: visualizing thematic changes in large document collections. *IEEE Transactions on Visualization and Computer Graphics* 8, 1, 9-20. DOI:10.1109/2945.981848.
- [22] W. Dou, X. Wang, R. Chang, and W. Ribarsky. 2011. ParallelTopics: A probabilistic approach to exploring document collections. In *2nd IEEE Conference on Visual Analytics Science and Technology 2011, VAST 2011*, Providence, RI, 231-240. DOI:10.1109/VAST.2011.6102461.
- [23] S. F. Liu, M. X. Zhou, S. Pan, W. Qian, W. Cai, and X. Lian. 2009. Interactive, topic-based visual text summarization and analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management, CIKM '09, Hong Kong, China — November 02 - 06, 2009* ACM, New York, NY, USA, 543-552. DOI:10.1145/1645953.1646023.
- [24] W. Ribarsky, D. X. Wang, and W. Dou. 2014. Social media analytics for competitive advantage. *Computers and Graphics* 38, 1, 328-331. DOI:<http://dx.doi.org/10.1016/j.cag.2013.11.003>.
- [25] S. Fortunato. 2010. Community detection in graphs. *Physics reports* 486, 3-5, 75-174. DOI:<https://doi.org/10.1016/j.physrep.2009.11.002>.
- [26] S. Wasserman and K. Faust. 1994. *Social network analysis: Methods and applications*. Cambridge University Press, Cambridge, UK.
- [27] J. Heinrich, D. Weiskopf, and 95-116. . . . In Eurographics (Stars). 2013. State of the Art of Parallel Coordinates. In *Eurographics 2013 - State of the Art Reports*, M. Sbert and L. Szirmay-Kalos Eds. The Eurographics Association, 95-116. DOI:<http://dx.doi.org/10.2312/conf/EG2013/stars/095-116>.
- [28] A. Inselberg and B. Dimsdale. 1987. Parallel Coordinates for Visualizing Multi-Dimensional Geometry. In *Computer Graphics 1987.*, T.L. Kunii Ed. Springer, Tokyo. DOI:10.1007/978-4-431-68057-4_3.
- [29] R. Zafarani and H. Liu. 2009. Social computing data repository at ASU. <http://socialcomputing.asu.edu/>
- [30] J. Leskovec and A. Krev. 2014. SNAP Datasets: Stanford Large Network Dataset Collection. <https://snap.stanford.edu/data/>

- [31] J. Yee, R. F. Mills, G. L. Peterson, and S. E. Bartczak. 2005. *Automatic Generation of Social Network Data from Electronic-Mail Communications*. Defense Technical Information Center.
- [32] H. Pérez-Rosés and F. Sebé. 2015. Synthetic generation of social network data with endorsements. *Journal of Simulation* 9, 4, 279-286. DOI:<https://doi.org/10.1057/jos.2014.29>.
- [33] A. Prat and X. Sánchez. 2017. Ldbc-snb data generator. https://github.com/ldbc/ldbc_snb_datagen
- [34] O. Erling, A. Averbuch, J. Larriba-Pey, H. Chafi, A. Gubichev, A. Prat, M.-D. Pham, and P. Boncz. 2015. The LDBC Social Network Benchmark: Interactive Workload. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data (SIGMOD '15)*. Melbourne, Victoria, Australia — May 31 - June 04, 2015 ACM, New York, NY, USA, 619-630. DOI:10.1145/2723372.2742786.
- [35] M.-D. Pham, P. Boncz, and O. Erling. 2012. S3G2: A Scalable Structure-Correlated Social Graph Generator. In *Selected Topics in Performance Evaluation and Benchmarking. TPCTC 2012*, R. Nambiar and M. Poess Eds. Springer, Berlin, Heidelberg. DOI:10.1007/978-3-642-36727-4_11.
- [36] A. Lenhart. 2015. *Teen, Social Media and Technology Overview 2015. Smartphones facilitate shifts in communication landscape for teens*. Pew Research Center.
- [37] S. Greenwood, A. Perrin, and M. Duggan. 2016. *Social Media Update 2016. Facebook usage and engagement is on the rise, while adoption of other platforms holds steady*. Pew Research Center.
- [38] B. Shneiderman. 1996. The eyes have it: A task by data type taxonomy for information visualizations. In *Proceedings of 1996 IEEE Symposium on Visual Languages. (3-6 Sept. 1996, Boulder, CO, USA, USA)* IEEE, EEUU, 336-343. DOI:10.1109/VL.1996.545307.
- [39] D. A. Keim, F. Mansmann, and J. Thomas. 2010. Visual analytics: how much visualization and how much analytics? *ACM SIGKDD Explorations Newsletter* 11, 2, 5-8. DOI:10.1145/1809400.1809403.
- [40] J. C. Roberts. 2007. State of the Art: Coordinated & Multiple Views in Exploratory Visualization. In *Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007)*, 61-71. DOI:10.1109/CMV.2007.20.
- [41] J. Nielsen and T. K. Landauer. 1993. A mathematical model of the finding of usability problems. In *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems. Amsterdam, The Netherlands — April 24 - 29, 1993* ACM, New York, NY, USA, 206-213. DOI:10.1145/169059.169166.
- [42] W. A. Pike, J. Stasko, R. Chang, and T. A. O'connell. 2009. The science of interaction. *Information Visualization* 8, 4, 263-274. DOI:10.1057/ivs.2009.22.
- [43] J. Durán-Escudero, F. J. García-Peñalvo, and R. Therón-Sánchez. An architectural proposal to explore the data of a private community through visual analytic. In J. M. Dodero, M. S. Ibarra Saiz, & I. Ruiz Rube (Eds.), *Proceedings of the Fifth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'17) (Cádiz, Spain, October 18-20, 2017)* (Article 48). ACM, New York, NY, USA. DOI:10.1145/3144826.3145398