

# Turning the TAP on Writing Analytics

**Antonette Shibani**

University of Technology Sydney  
antonette.shibani@gmail.com

**Sophie Abel**

University of Technology Sydney  
sophie.abel@student.uts.edu.au

**Andrew Gibson**

University of Technology Sydney/ Queensland University of Technology  
andrew.gibson@qut.edu.au

**Simon Knight**

University of Technology Sydney  
sjgknight@gmail.com

**ABSTRACT:** Writing analytics is seen as a potentially useful technique that uses textual features to provide formative feedback on students' writing. However, for this feedback to be effective, it is important that it is aligned to pedagogic contexts. Such efficient integration of technology in pedagogy could be supported by developing writing analytics literacy. The proposed workshop aims to build this capacity by mapping technical constructs to a broader educational sense for pragmatic applications. It provides a hands-on experience for participants to work with text analytics and discuss its implications for writing feedback. Participants will work with a set of text analysis code to extract features and map them to writing feedback. They will also be given opportunity to develop rules based on extracted text features to write feedback for their own pedagogic contexts.

**Keywords:** Writing analytics, learning analytics, text mining, writing, writing analytics literacy, data carpentry, hackathon

## 1 BACKGROUND

Across educational contexts students' written communication is a fundamental concern (National Commission On Writing, 2003; OECD, 2013). Research and commercial tools built on Natural Language Processing (NLP) technologies have gained some traction in supporting the analysis of this writing for educational purposes (e.g., McNamara, Graesser, McCarthy, & Cai, 2014; Shermis & Burstein, 2013). In learning analytics, the sub-domain of *writing analytics* has emerged to support students in their writing practices through the provision of formative feedback that provides insights to educators and students. Two previous workshops (Buckingham Shum et al., 2016; Knight, Allen, Gibson, McNamara, & Buckingham Shum, 2017) have focused on this topic, with community building events focusing on critical perspectives on writing analytics (Buckingham Shum et al., 2016), and the development of greater 'literacy' for writing analytics – building capacity for the use of analytics through its alignment with pedagogic ends (Knight et al., 2017).

This proposed workshop, which would be the third in the series, will build on these previous events by introducing participants to a tool: the Text Analytics Pipeline (TAP). The workshop will build a shared approach to mapping low level textual features to rules that are based on empirical and theoretical work, which can be translated into feedback for students and educators. The LAK17 workshop focused on a need to go beyond simple text analytics, to think about how educators might make effective use of the power of NLP in their teaching. In that workshop the focus was on how a set of existing tools, submitted by participants, map to particular problems in writing, and how each of those tools provide feedback. In the current workshop, participants will work hands-on with text features by generating them using sample code in notebooks, and mapping these features to pedagogic contexts alongside writing useful feedback drawing on the features. At a simple level this might involve, for example, understanding how basic named entity recognition features can be used to develop rules to provide feedback to students: “You’ve included x, y, and z, but isn’t v also an important researcher in this area?”. Discussion will focus on technical concerns, and the pedagogic, drawing on research in the pedagogy of teaching writing, and the potential and pitfalls of NLP in addressing that pedagogy. This aligns closely with the theme of user-centered analytics by involving different stakeholders in the design of writing analytics feedback.

For educators to make effective use of writing analytics tools for impact on learning, tools must be integrated into teaching and learning contexts where they guide action, connecting theory, pedagogy, and assessment (Clow, 2012; Knight, Buckingham Shum, & Littleton, 2014; Shibani, Knight, Buckingham Shum, & Ryan, 2017; Wise & Shaffer, 2015; Wise, Vytasek, Hausknecht, & Zhao, 2016). Thus, the third workshop is intended to:

1. Build synergy between writing analytics literacy and writing assessment literacy – that is, build understanding both of the potential of writing analytics and of how to assess writing (both using, and without, analytics tools)
2. Build practitioner capacity for research on their students writing, through developing understanding of how data from writing analytics might provide insights on that writing
3. Build student writing analytics literacy as a means to develop their writing via their critical interaction with text features that contribute to good writing

This workshop will adopt a data carpentry approach to teach writing analytics constructs. Data Carpentry workshops teach beginners the foundation skills needed to conduct data driven research. They are based on Software Carpentry bootcamps (Wilson, 2006), where beginners are taught “basic concepts, skills, and tools for working with data so researchers can get more done in less time and with less pain” (Teal et al., 2015, p. 1). As learners come with varied prior experience, the sessions are an opportunity to build their toolkit so that they can start working with data in their own research. This workshop will guide participants through some vignettes that illustrate the use of a tool (TAP), and its connection to pedagogy, and feedback. The workshop is thus not solely *technical* in nature, instead focusing on developing shared alignment between technical, and social features, towards feedback delivered through analytics tools.

## 2 SUBMISSIONS AND WORKSHOP FORMAT

### 2.1 Workshop Objectives

The workshop aims to build writing analytics capacity through developing writing analytics literacy both at this event and beyond. Workshop activities will have two foci: the technical, and the social, both targeted at provision of feedback that supports student writing. The workshop will engage participants with the Text Analytics Pipeline (TAP), and the mapping of textual features output by TAP to rules that can be used to provide feedback to students. The connection between the technical and the pedagogic will be the focus of hands-on activity and will also frame discussion.

### 2.2 Workshop Activities and Half-day Schedule

This half day workshop will take a participatory approach, blending workshop, tutorial, and hackathon to consider the potential of text analytics tools for supporting writing, and how through use of openly available tools and a data carpentry approach, novices to text analytics – including educators, learning technologists, and others – can be inducted into its potential. The tentative schedule is given below:

**Introduction (15 min):** Introducing the objectives, presenters and technical set-up for participants.

**Introduction to the notebook/workbook (20 mins):** The basics of using Jupyter notebooks with some simple text analysis examples.

**Tutorial and discussion Part I (50 min):** i) Hands-on engagement with text analysis using Jupyter notebooks – extracting features from text.

**Break (15 min)**

**Tutorial and discussion Part II (60 min):** i) Hands-on engagement with text analytics – turning features into feedback.

**Open discussion (30 min):** Discussion and co-creation of a shared resource that maps how other text features can be used in pedagogic applications.

**Closing remarks (15 min):** Brief summary and discussion of the workshop and future steps.

### 2.3 Participation, Required Equipment and Dissemination

Participation will be 'open' (i.e., any interested delegate may register to attend). The workshop does not require any special equipment (wifi and a room with power strips aside). Participants are encouraged to bring their own devices (laptops best or tablets with keyboards) with a modern web browser. We expect 15-30 participants to attend. An invitation will be extended to participants of previous Writing Analytics workshops to bring different perspectives on the textual features that can be identified and the kinds of feedback that can be provided to help students improve their writing. This workshop will be of interest to a wide range of LAK delegates including: students and researchers actively engaged in writing research, text analytics or writing analytics specifically; educators in schools, universities and businesses; leaders and policymakers; and companies active or potentially active in the field. Some coding skills although not mandatory might be useful.

The workshop organizers are embedded in the learning analytics and related communities. They will make use of listservs (SoLAR, Learning Analytics Google group, EDM-announce, ISLS, SIG-LS, EARLI, ICCE, CHI) and leverage their own personal networks to advertise the workshop. Researchers, practitioners, and funders indicate an increasing interest in writing analytics, and approaches to put

writing analytics into practice are currently at the forefront of many learning analytics efforts, thus we anticipate the workshop having popular appeal.

## REFERENCES

- Buckingham Shum, S., Knight, S., McNamara, D., Allen, L. K., Betik, D., & Crossley, S. (2016). Critical Perspectives on Writing Analytics (pp. 481–483). Presented at the 6th ACM Learning Analytics and Knowledge Conference, Edinburgh, UK: ACM. <https://doi.org/10.1145/2883851.2883854>
- Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134–138). Retrieved from <http://dl.acm.org/citation.cfm?id=2330636>
- Knight, S., Allen, L. K., Gibson, A., McNamara, D., & Buckingham Shum, S. (2017). Writing analytics Literacy - Bridging from Research to Practice. Presented at the 7th ACM Learning Analytics and Knowledge Conference, Edinburgh, UK: ACM.
- Knight, S., Buckingham Shum, S., & Littleton, K. (2014). Epistemology, assessment, pedagogy: where learning meets analytics in the middle space. *Journal of Learning Analytics*, 1(2). Retrieved from <http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/3538>
- McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge, UK: Cambridge University Press.
- National Commission On Writing. (2003). *Report of the National Commission on Writing in America's Schools and Colleges: The Neglected "R," The Need for a Writing Revolution*. College Board. Retrieved from [http://www.collegeboard.com/prod\\_downloads/writingcom/neglectedr.pdf](http://www.collegeboard.com/prod_downloads/writingcom/neglectedr.pdf)
- OECD. (2013). PISA 2015: Draft reading literacy framework. OECD Publishing. Retrieved from <http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Reading%20Framework%20.pdf>
- Shermis, M. D., & Burstein, J. (2013). *Handbook of Automated Essay Evaluation: Current Applications and New Directions*. Routledge.
- Shibani, A., Knight, S., Buckingham Shum, S., & Ryan, P. (2017). Design and Implementation of a Pedagogic Intervention Using Writing Analytics. In W. Chen et al. (Ed.). *Proceedings of the 25th International Conference on Computers in Education*, New Zealand.
- Teal, T. K., Cranston, K. A., Lapp, H., White, E., Wilson, G., Ram, K., & Pawlik, A. (2015). Data carpentry: workshops to increase data literacy for researchers. *International Journal of Digital Curation*, 10(1), 135–143. Retrieved from <http://ateson.net/d/49962/index.php/ijdc/article/download/10.1.135/386>
- Wilson, G. (2006). Software carpentry. *Computing in Science & Engineering*, 8, 66. Retrieved from <http://scitation.aip.org/content/aip/journal/cise/8/6/10.1109/MCSE.2006.122?crawler=true>
- Wise, A. F., & Shaffer, D. W. (2015). Why Theory Matters More than Ever in the Age of Big Data. *Journal of Learning Analytics*, 2(2), 5–13. Retrieved from <http://epress.lib.uts.edu.au/journals/index.php/JLA/article/view/4677>
- Wise, A. F., Vytasek, J. M., Hausknecht, S., & Zhao, Y. (2016). Developing Learning Analytics Design Knowledge in the "Middle Space": The Student Tuning Model and Align Design Framework for Learning Analytics Use. *Online Learning*, 20(2). Retrieved from <http://olj.onlinelearningconsortium.org/index.php/olj/article/view/783>