Learning Analytics as a "Middle Space"

Dan Suthers
Dept. of Information and Computer Sciences,
University of Hawaii at Manoa
1680 East West Road, POST 309
Honolulu, HI 96822 USA
suthers@hawaii.edu

Katrien Verbert
Department of Computer Science, Eindhoven
University of Technology
P.O. Box 513 5600 MB Eindhoven, The
Netherlands
k.verbert@tue.nl

ABSTRACT

Learning Analytics, an emerging field concerned with analyzing the vast data "given off" by learners in technology supported settings to inform educational theory and practice, has from its inception taken a multidisciplinary approach that integrates studies of learning with technological capabilities. In this introduction to the Proceedings of the Third International Learning Analytics & Knowledge Conference, we discuss how Learning Analytics must function in the "middle space" where learning and analytic concerns meet. Dialogue in this middle space involves diverse stakeholders from multiple disciplines with various conceptions of the agency and nature of learning. We hold that a singularly unified field is not possible nor even desirable if we are to leverage the potential of this diversity, but progress is possible if we support "productive multivocality" between the diverse voices involved, facilitated by appropriate use of boundary objects. We summarize the submitted papers and contents of these Proceedings to characterize the voices and topics involved in the multivocal discourse of Learning Analytics.

Categories and Subject Descriptors

K.3 [**Computers and Education**]: General

Keywords

learning analytics, multidisciplinarity, boundary objects, productive multivocality

1. DIALECTICS IN LAK

The International Learning Analytics & Knowledge (LAK) Conference, now in its third year[1], is a venue for reporting and advancing research that is motivated by the nexus of two emerging societal phenomena. First we are witnessing rapidly expanding use of technologies in supporting learning, not only in established institutional contexts and platforms, but also in online settings offering free and open learning opportunities. Second, the unprecedented availability of data that learners generate in the process of accessing materials, interacting with educators and peers, and creating new content in these technological settings, coupled with advances in analytics and data mining, knowledge modeling, and open data offer great potential for research into how learning takes place in socio-technical settings and the development of new forms of analytics that can inform learners and educators. Learning Analytics research seeks to bring these technical, pedagogical, and social domains into dialogue with each other to ensure that interventions and organizational systems serve the needs of all stakeholders [7].

The third year of a new conference series might be seen as a critical year in which the identity of an emerging field of study begins to stabilize. Learning Analytics from its inception has sought to take a multidisciplinary approach to integrating studies of learning with technological capabilities. LAK 2011 called for a "focus on integrating the technical and the social/pedagogical dimensions of learning analytics"[2]. LAK 2012 was advertised as "a multidisciplinary conference for: learning scientists; (computer) scientists and data/knowledge engineers; researchers in education, sociology, psychology, information science; educators at all levels; Institutional/Organizational data analysts; training and development professionals; educational and academic leaders; business leaders; [and] course management system developers and leaders"[3]. At that conference, papers addressed diverse educational issues and approaches to those issues that leverage the availability of data about learning with computational, representational and visualization techniques, but the first author concluded that learning and analytic concerns needed to be integrated more tightly.

This year, we sought to address these needs and bring these many voices into dialogue under the overarching theme of "Dialectics in Learning Analytics". This theme has three facets. First, to explore the *Middle Space* within which Learning and Analytics intersect, we called for papers and events that explicitly connect analytic tools to theoretical and practical aspects of understanding and managing learning. Second, we planned the program to foster *Productive Multivocality* between the diverse "voices" or perspectives that come from different disciplines, theories, research methods, and roles within the educational enterprise. Third,

[1] http://lakconference2013.wordpress.com/

[2] https://tekri.athabascau.ca/analytics/

[3] http://lak12.sites.olt.ubc.ca/

LAK '13, April 08 - 12 2013, Leuven, Belgium

there is a dialectic between *the Old and the New*: we are facing the centuries-old problem of improving learning, but bringing established knowledge together with a new set of tools not available before. Also, we address these problems in the city of Leuven: centuries old, lively new. Below we elaborate on the first two themes.

2. THE MIDDLE SPACE

Our advice to authors[4] requested that papers address the "middle space" where learning and analytics meet. Some papers may make contributions primarily to analytic technologies, while others may contribute primarily to learning theory or educational practice, but learning analytics calls for work that address each in relation to the other.

Learning analytics is about learning. Learning analytics research should be explicit about the theory or conception of learning underlying the work, and manifest this conception in the work presented. We hold that no particular theory or conception of learning should be favored a-priori: individuals, small groups, and/or larger collectives may be the agent of learning; and learning may consist of knowledge or skill acquisition, intersubjective meaning-making, or changes in identity and participation in a community, among other processes [9]. Furthermore, learning can be conceived of as simultaneously taking place at all of these granularities of agency and involving all of these epistemological processes. Research that analyzes learning processes across multiple granularities and brings multiple methodological and theoretical orientations to bear are particularly appropriate for understanding learning as a complex phenomenon. Regardless of how learning is conceptualized, the point is to encourage learning analytics researchers to make their conception explicit and reflect on how their analytic approach relates to their conception of learning, so that we can be mindful about our assumptions in dialogue with each other.

Learning analytics is about analytics. Learning analytics research should be explicit about how the work either offers relevant new analytic methods (e.g., computational, representational, statistical, and visualization methods) or advances our understanding of the value of existing analytic methods, in either case for understanding learning and educational practices. Research on learning analytics may vary in the degree to which it makes technical contributions, but the connection to learning should be present. Research making a strong technical contribution need not also include a study in an applied setting, but should at least discuss how the properties of the technical contribution are relevant to understanding or managing learning in practice. Research with a learning theoretic or practical contribution need not advance the technical state of analytics, but should at least discuss or evaluate whether and how the affordances of the chosen analytic methods bring forth relevant aspects of the data in a manner enabling the primary contribution. Thus, we hope that the field will avoid simplistic judgments about work being "merely technical" vs. "not innovative" that are sometimes heard, but rather bring these kinds of contributions into coordination.

In summary, although individual research efforts may differ in their emphasis, we believe that all research in Learning Analytics should address the "middle space" by including both learning and analytic concerns and addressing the match between technique and application. Advances in learning theory and practice are welcome, provided that they are accompanied with an evaluation of how existing or new analytic technologies support such advances. Advances in analytic technologies and methods are welcome, provided that they are accompanied with an evaluation of how understanding of learning and educational practices may be advanced by such methods.

[4]http://lakconference2013.wordpress.com/for-authors/

3. PRODUCTIVE MULTIVOCALITY

Learning analytics is multidisciplinary, drawing on theories and methods from diverse research traditions. Our community includes educators, learning scientists, computer scientists, administrators, and policy makers, among others (see also the LAK 2012 call quoted above). Many "voices" are brought together, leading to the question of how such multivocality can be productive. Is a unified field of "Learning Analytics" possible?

The situation is similar to that of the Computer Supported Collaborative Learning research community, wherein efforts have been made to bridge across theoretical and methodological traditions [5, 10], leading to the following observations that apply to Learning Analytics as well. The field is too diverse (in theory, methods, and even objectives) for unification to be possible; nor would it be desirable, as this diversity is a potential strength of Learning Analytics. However, this strength can only be realized if these voices are brought into dialogue (or multi-logue) with each other, avoiding balkanization into disjoint dialogues nominally under "learning analytics and knowledge". This tension between the impossibility of unification and the need to dialogue might be served by *boundary objects* [8], physical or conceptual entities that each tradition interprets in its own way, but that provide common referents or points of articulation to ground conversations.

Boundary objects can exist and be leveraged in multiple layers. The middle space itself serves as a topical boundary object, as Learning Analytics brings multidisciplinary voices into discourse around the question of how analytic tools can help our understanding of learning and design of educational practices. A focus on data grounds discussion in concrete situations where differing conceptions are uncovered in their implications for learning and educational practice. Yet, shared data is not enough: members of different traditions may "talk past" each other, construing the data in entirely different ways and addressing incommensurable concerns. An attempt at shared analytic objectives is helpful: even if the objectives are interpreted in different ways, this provides a second dimension along which assumptions can be exposed and compared. Analytic representations such as visualizations that are meaningful for members of multiple research traditions or for both technologists and practitioners can play important roles in this process. Bringing representations from different analytic sources into alignment into each other highlights concordances and discordances. Visual representations are accessible in some manner to technologists and practitioners alike, and can be a focus for questions about whether analytic computations are exposing features that are pedagogically *actionable*. The middle space is where we also find boundary objects in the form of specific pedagogical objectives or interventions, some of which are noted below.

4. LEARNING ANALYTICS AS REFLECTED IN THESE PROCEEDINGS

Thus far we have been speaking of ideals, while a research community is also defined by the practices of its participants. Who is participating in the Learning Analytics dialogue, and what topics do we find in this middle space? These questions might best be answered on an empirical basis, as reflected in the present proceedings.

Submissions were received with author affiliations from 31 countries world-wide, and accepted papers include affiliations from Australia, Belgium, Canada, the Czech Republic, Denmark, Ecuador, France, Germany, Greece, Hong Kong, Italy, Malta, the Netherlands, Norway, Spain, South Africa, Switzerland, the United Kingdom, and the United States of America. There were 58 full paper submissions, 37 short paper submissions (including 2 classified as design briefings), and also various panel, workshop and tutorial proposals. Each paper was reviewed by at least three reviewers, and our decisions were based on the content of the reviews as well as numerical rankings. Of the full papers, 16 (27.6%) were accepted as such, and 14 (24.1%) as short papers (with one conversion to a design briefing). Of the short papers, 8 (22.9%) were accepted in that category. Some other submissions not accepted as papers were encouraged to resubmit as posters, and 11 were accepted as such. We also accepted or solicited 3 panels, and the Workshop and Tutorial chairs accepted 5 of 10 workshop proposals and solicited 3 excellent tutorials.

Let us examine how the topics of these papers characterize the field of Learning Analytics. The following themes are emerging in the learning analytics area (see also [7, 11]):

Reflections on Learning Analytics. This section of the proceedings includes the present extended abstract and a contribution by Balacheff and Lund that addresses the conference theme, going into depth on the multi-disciplinarity of LAK and its potential for productive miltivocality.

Visualizations for reflection and awareness. Several contributions focus on increasing reflection and awareness through the use of visualization techniques. The focus is on collecting traces that learners leave behind and visualizing those traces to improve learning [2]. Dashboard applications are presented as well as evaluation studies that assess the impact of such dashboards on learning. Such dashboards are focusing on the analysis and visualization of different learning indicators to foster awareness and reflection about learning processes [1]. These indicators include resource accesses, time allocated, and knowledge level indicators. Visualization methods appear throughout the other categories, reflecting their relevance to multiple stakeholders.

Social network analysis and visualization. Social learning analytics [3] is the core research topic in three collections of contributions. In the first, analysis and visualization of social interactions is researched to make people aware of their social context and to enable them to explore this context [4]. In TEL, this is particularly, but not only, relevant for Computer Supported Collaborative Learning (CSCL), where the interactions with peer learners are a core aspect of how learning is organized.

Communication and collaboration. Papers in this second social learning area focus on both reading and writing activities, including analysis of paragraph level revisions and how collaboration forms around topics in collaborative writing, analyses of discussion forums that includes temporal patterns of reading as well as writing and how analytics can support learner self-regulation, and how students' promotion of each others' work in a blogging environment can be leveraged for feedback.

Discourse analytics. Papers in this section have some affinity to those in Sequence Analysis further on as well to social learning analytics. Discourse is approached from diverse perspectives, including scientometric path analysis (citation analysis that includes temporality) to model knowledge evolution in Wikiversity; a comparative evaluation of seven automatic classification approaches for identifying exploratory dialogue; and theoretical reflections on how learning analytics embody assumptions about both our own and learners' epistemologies, with applications to analysis of user trace data.

Behavior analysis. Papers classified here offer advances in making higher level inferences from low level actions. The papers generally start with fine-grained behavior, but are wide ranging in the nature of the data and scope of analysis. Three papers reflect the importance of extending our scope of analysis to include multimodal behavior such as gesture and object manipulation, and using alternative sources of data such as eye tracking, in this case to analyze joint attention. A fourth addresses how to uncover higher leverl learning processes based on low level data in the Kahn Academy.

Affect analytics. Affect is discussed in two full paper contributions, including novel ideas to detect emotional aspects from eye-tracking data. Eye-tracking data is also used by several other contributions to detect behavior patterns. Such patterns are used to estimate confidence or the level of expertise.

Predictive analytics. Prediction of learner performance and modeling of learners have been researched extensively by the educational data mining, educational user modeling and educational adaptive hypermedia communities. The objective is to estimate the unknown value of a variable that describes the learner, such as performance, knowledge, scores or learner grades [6]. Predictive analytics that examine extraction of learning indicators are researched by two contributions–among others to learn from Facebook data and to predict at-risk students.

Sequence analysis. The temporal sequencing of activity contains useful information about learners. While concerns with sequentiality also surface elsewhere (e.g., in communication and collaboration, and discourse analytics), two papers are gathered here for their explicit focus on sequential structure of behavior. They do so at two very different granularities: microgenetic sequence analysis of the development of student understanding of fractions in a game environment, and sequences of course enrollment and completion as a student works through a self-designed program of study.

MOOCs. The proceedings shifts towards an applications focus, beginning with learning analytics as applied to Massive Open Online Courses (MOOCs) in which large numbers of people participate with an open enrollment model. Although enrollments can be very high, completion rates are lower than in traditional courses, giving urgency to research on motivations for participation and understanding how learners may obtain value from a MOOC under different participation trajectories, even if they do not formally complete the course.

Assessment. The objective of 4 contributions is to use data that is tracked from the actual use of learning environments to support assessment processes. The papers address the generalizability of data-mining assessment models, and assessments based on rubrics in an LMS, on student interaction with online lectures, and on models of the cognitive processes involved in a learning task.

Supporting Teachers. Assessment is a natural lead-in to work that is explicitly concerned with support for educators. Three contributions focus specifically on teacher support– among others by visualizing data collected in real applications to monitor and direct content development. Work conducted in collaboration with teachers has great potential for creative synergies in the middle space.

Challenges with respect to scalable learning analytics, ethics and institutional policies for using student data, and collection and sharing of datasets for research purposes are discussed by three contributions.

Analytic Architectures. Two papers offering architectures for managing learning analytics data address some of the above challenges with frameworks for organizing the analytic enterprise that support analysis across micro and macro levels of activity, and across the various tools used in a collaborative learning environment.

Design Briefings. The final set of papers explore analytics in the context of specific system designs, including how analytics can support gamification of learning management systems, a system that analyzes classroom video in real time to automatically assess levels of student attention, and examination of three ways in which learning analytics can support orchestrate community inquiry in the SAIL environment.

Panels and Workshops. The proceedings ends with abstracts for panels and workshops to be held at the conference. Panels examine institutional and individual engagement with learning analytics; while workshops explore new discourse analytic approaches, and the specific analytic challenges offered by video-based learning, learning resource repositories, and teaching as objects of study.

Thus we have in this proceedings a rich collection of analytic methods, and of learning settings and educational practices that the methods might inform. We are well poised for discussion that keeps learning and analytics simultaneously in focus. By being explicit about our assumptions and leveraging multiple forms of boundary objects, we can embark on productively multivocal discourse that leverages the diversity of the Learning Analytics and Knowledge community.

5. ACKNOWLEDGMENTS

We thank Erik Duval and Xavier Ochoa for their support and discussion of the conference theme, and the reviewers for their help in providing high quality reviews of submitted contributions. Discussion with Janet Kolodner also influenced the "middle space" theme. Katrien Verbert is a post-doctoral fellow of the Research Foundation - Flanders (FWO).

6. REFERENCES

[1] E. Duval, J. Klerkx, K. Verbert, T. Nagel, S. Govaerts, G. A. Parra Chico, J. L. Santos Odriozola, and B. Vandeputte. Learning dashboards & learnscapes. In *Workshop at CHI2012: ACM SIGCHI Conference on Human Factors in Computing Systems*, May 2012.

[2] E. Duval and K. Verbert. Learning analytics. *eleed*, 8, 2012.

[3] R. Ferguson and S. B. Shum. Social learning analytics: five approaches. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, LAK '12, pages 23–33, New York, NY, USA, 2012. ACM.

[4] J. Heer and D. Boyd. Vizster: Visualizing online social networks. In *Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization*, INFOVIS '05, pages 5–, Washington, DC, USA, 2005. IEEE Computer Society.

[5] T. Koschmann, editor. *Theories of Learning and Studies of Instructional Practice*. New York: Springer, 2011.

[6] C. Romero and S. Ventura. Educational data mining: A survey from 1995 to 2005. *Expert Syst. Appl.*, 33(1):135–146, July 2007.

[7] G. Siemens. Learning analytics: envisioning a research discipline and a domain of practice. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, LAK '12, pages 4–8, New York, NY, USA, 2012. ACM.

[8] S. L. Star and J. R. Griesemer. Institutional ecology, 'translations' and boundary objects: Amateurs and professionals in berkeley's museum of vertebrate zoology, 1907-39. *Social Studies Of Science*, 19(3):387–420, 1989.

[9] D. D. Suthers. Technology affordances for intersubjective meaning making: A research agenda for CSCL. *The International Journal of Computer-Supported Collaborative Learning*, 1(3):315–337, 2006.

[10] D. D. Suthers, K. Lund, C. P. Rosé, G. Dyke, N. Law, C. Teplovs, W. Chen, M.M. Chiu, H. Jeong, C-K. Looi, R. Medina, J. Oshima, K. Sawyer, H. Shirouzu, J-W. Strijbos, S. Trausan-Matu, J. van Aalst. Towards productive multivocality in the analysis of collaborative learning. In *Connecting Computer-Supported Collaborative Learning to Policy and Practice: Proceedings of the 9th International Conference on Computer-Supported Collaborative Learning (CSCL 2011) (Vol. III)*, pages 1015–1022. Hong Kong: International Society of the Learning Sciences, 2011.

[11] K. Verbert, N. Manouselis, H. Drachsler, and E. Duval. Dataset-driven research to support learning and knowledge analytics. *Educational Technology and Society*, 15(3):133–148, June 2012.