

Current and Future Multimodal Learning Analytics Data Challenges

Daniel Spikol
Malmö University
Malmö, Sweden
daniel.spikol@mah.se

Marcelo Worsley
Northwestern University
Evanston, IL, USA
mw@northwestern.edu

Luis P. Prieto
Tallinn University
Tallinn, Estonia
lprisan@tlu.ee

Xavier Ochoa
ESPOL
Guayaquil, Ecuador
xavier@cti.espol.edu.ec

M. J. Rodríguez-Triana
REACT Group, EPFL
Lausanne, Switzerland
maria.rodrigueztriana@epfl.ch

Mutlu Cukurova
UCL Knowledge Lab
London, UK
m.cukurova@ucl.ac.uk

ABSTRACT

Multimodal Learning Analytics (MMLA) captures, integrates and analyzes learning traces from different sources in order to obtain a more holistic understanding of the learning process, wherever it happens. MMLA leverages the increasingly widespread availability of diverse sensors, high-frequency data collection technologies and sophisticated machine learning and artificial intelligence techniques. The aim of this workshop is twofold: first, to expose participants to, and develop, different multimodal datasets that reflect how MMLA can bring new insights and opportunities to investigate complex learning processes and environments; second, to collaboratively identify a set of grand challenges for further MMLA research, built upon the foundations of previous workshops on the topic.

CCS Concepts

• **Applied computing** → **Interactive learning environments**; *Collaborative learning*; • **Information systems** → *Data analytics*;

Keywords

Multimodal learning analytics; datasets; challenges.

1. WORKSHOP BACKGROUND

The field of learning analytics is shifting from an emerging area of research, into mainstream educational research and practice. The inherently blended nature of most current learning settings, makes it essential to move beyond learning analytics that rely solely on a single data source (usually log files). Multimodal learning analytics (MMLA) provides insights into such learning processes that happen across multiple contexts between people, devices and resources (both

physical and digital), which often are hard to model and orchestrate [4, 5, 7, 8].

MMLA leverages the increasingly widespread availability of sensors and high-frequency data collection technologies to enrich the existing data available. Using such technologies in combination with machine learning and artificial intelligence techniques, LA researchers can now perform text, speech, handwriting, sketch, gesture, affective, neurophysical, or eye gaze analyses [2, 6].

The workshop builds upon previous MMLA workshops at LAK, focusing on hands-on exposure to multimodal datasets and techniques for analysis. In addition, this workshop introduced the opportunity for participants to present their own multimodal datasets, as well as future-looking MMLA challenges and concepts, helping define this emerging field through discussion and feedback.

1.1 MMLA Community Foci

This full-day workshop was organised following the two main emphases of this emergent community. The first one is to build up and expand the community by letting participants explore multimodal techniques and data, using pre-made environments and sample datasets. This helps newcomers to learn and explore, and lets experienced participants see and discuss different approaches. Furthermore, in this workshop participants also were encouraged to submit contributions proposing new MMLA datasets that the community can use in the future, as well as future-oriented visions and challenges.

The discussion on the initial and participant-submitted multimodal datasets, and the techniques that can be used to gather and analyze them, set the stage for the second emphasis of the community: unearthing the major future challenges in MMLA research, and steering the community towards overcoming them.

2. WORKSHOP OBJECTIVES

The workshop has two interrelated objectives that frame the emerging interest in MMLA. The workshop aims to create a ‘common ground’ for better understanding the current state of research and practice. As the MMLA community grows, it is also necessary to launch a set of future grand challenges to guide research, practice, and to bring together the different viewpoints of the field.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

LAK '16 Vancouver, BC Canada

© 2017 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-4870-6/17/03.

DOI: <http://dx.doi.org/10.1145/3027385.3029437>

Disseminate the state of MMLA research through sharing datasets: Building upon previous workshop editions (in the LAK and ICMI conferences), this workshop explored advanced approaches and techniques for working with multimodal datasets. By providing access to high-quality datasets, participants were able to contrast analyses, triggering in-depth discussions and learning about the kinds of insights that can be gained from different techniques.

Discuss the current and future MMLA challenges: By letting participants submit their own multimodal datasets, this workshop continued previous efforts, expanding the discussion about what constitutes good practice for the design of data and analysis of MMLA. Furthermore, as the MMLA community grows, it was necessary to bring together different viewpoints of the field, to help identify future challenges, and define an agenda that guides MMLA in advancing research and practice.

2.1 Target Participants

The workshop welcomed researchers that submitted multimodal datasets, techniques or future challenges prior to the workshop. It also welcomed participants without a prior contribution, that desired to learn about MMLA and contribute to the discussion. All in all, participants came from different backgrounds:

Existing MMLA researchers. The LAK conference is the annual meeting for MMLA researchers, and continues to build upon the efforts of past workshops, accommodating the growing interest and work in MMLA.

New MMLA researchers. One of the key aims of the workshop is to attract new researchers from the LAK community to MMLA, and especially into its new Special Interest Group¹ (SIG). The low barrier to participate, with the hands-on session and the discussions, provided a dual opportunity to attract new researchers and practitioners.

LA researchers and practitioners. The workshop is also aimed at bringing together other LA researchers and practitioners that can discover there how to incorporate multimodal data and techniques into their ongoing research.

2.2 Initial Datasets

The workshop used the Math Data Corpus [3], Oral Presentation Quality [1], and PELARS Project² datasets. The first two data corpora have been used in previous workshops, and are well documented. On the other hand, the PELARS dataset was new and provided a contrast for discussion about what makes a good multimodal dataset. The participant-submitted datasets provided seeds and new possibilities for eventual use in future MMLA events.

2.3 Dissemination Strategy

The workshop had two main planned outputs to be disseminated. First, the sharing of existing techniques and datasets, created among the research community and interested stakeholders. Additionally, the participant submissions formed the foundation for a set of grand challenges and a research agenda for MMLA, to support the forthcoming MMLA-SIG within the SoLAR community. After the conference, full-length submissions resulting from the discussions will be published on CEUR, and we are exploring the joint edition of future journal special issues on the topic.

¹<http://sigmla.org/>

²<http://www.pelars.eu>

The workshop used the LAK17 channels for the call of participation. Additionally, the organizers used their network of contacts (and that of the Program Committee), to personally contact a broad range of researchers and practitioners that have overlapping interests in related fields. The new-born MMLA SIG will serve as the main channel to consolidate the interest around the workshop.

2.4 Program Committee

Apart from the workshop organisers, the following researchers will be part of the program committee: Alejandra Martínez Monés, UVA, Spain; Barbara Wasson, University of Bergen, Norway; Cynthia D'Angelo, SRI, USA; Davinia Hernández-Leo, UPF, Spain; Denise Whitelock, OU, UK; Manolis Mavrikis, UCL, UK; Nour El Mawas, Institut Mines-Télécom, France; Paulo Blikstein, Stanford, USA; Roberto Martinez-Maldonado, UTS, Australia.

3. ADDITIONAL AUTHORS

Additional authors: Bahtijar Vogel (Malmö University, bahtijar.vogel@mah.se), Emanuele Ruffaldi (Scuola Superiore Sant'Anna, e.ruffaldi@sssup.it), and Ulla Lunde Ringtved (University College Nordjylland, ulr@ucn.dk).

4. REFERENCES

- [1] L. Chen, G. Feng, J. Joe, C. W. Leong, C. Kitchen, and C. M. Lee. Towards automated assessment of public speaking skills using multimodal cues. In *Proceedings of the 16th International Conference on Multimodal Interaction*, ICMI'14, pages 200–203, New York, NY, USA, nov 2014. ACM Press.
- [2] P. J. Donnelly, N. Blanchard, B. Samei, A. M. Olney, X. Sun, B. Ward, S. Kelly, M. Nystran, and S. K. D'Mello. Automatic teacher modeling from live classroom audio. In *Proceedings of the 2016 Conference on User Modeling Adaptation and Personalization*, UMAP'16, pages 45–53, New York, NY, USA, jul 2016. ACM Press.
- [3] L.-P. Morency, S. Oviatt, S. Scherer, N. Weibel, and M. Worsley. ICMI 2013 Grand Challenge Workshop on Multimodal Learning Analytics. *ICMI'13*, pages 373–378, 2013.
- [4] X. Ochoa, M. Worsley, K. Chiluitza, and S. Luz. MLA'14: Third Multimodal Learning Analytics Workshop and Grand Challenges. *ICMI'14*, pages 531–532, 2014.
- [5] X. Ochoa, M. Worsley, N. Weibel, and S. Oviatt. Multimodal Learning Analytics Data Challenges. *LAK'16*, pages 498–500, 2016.
- [6] L. P. Prieto, K. Sharma, P. Dillenbourg, and M. J. Rodríguez-Triana. Teaching analytics: Towards automatic extraction of orchestration graphs using wearable sensors. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge - LAK'16*, pages 148–157, New York, NY, USA, apr 2016. ACM Press.
- [7] S. Scherer, M. Worsley, and L. Morency. 1st International Workshop on Multimodal Learning Analytics. *ICMI'12*, page 609, 2012.
- [8] M. Worsley, K. Chiluitza, J. F. Grafsgaard, and X. Ochoa. 2015 Multimodal Learning and Analytics Grand Challenge. *ICMI'15*, pages 525–529, 2015.