Automating Data Narratives in Learning Analytics Dashboards using GenAI

Adriano Pinargote[1,*], Eddy Calderón[1], Kevin Cevallos[1], Gladys Carrillo[1], Katherine Chiluiza[1] and Vanessa Echeverria[1,2]

[1]*Escuela Superior Politecnica del Litoral , (Information Technology Center), Campus Gustavo Galindo Km. 30.5 Vía Perimetral, P.O. Box 09-01-5863, Guayaquil, Ecuador*

[2]*Monash University, Clayton, VIC, Australia*

Abstract

This paper presents an innovative approach leveraging Generative Artificial Intelligence to automate data narratives within Learning Analytics Dashboards for collaborative learning scenarios. Focusing on the analysis of class meeting transcripts, the study delves into specific collaboration skill metrics, transforming raw data into a cohesive narrative. Validation through inter-rater reliability, utilizing Cohen's Kappa coefficient, establishes the reliability of both human and AI assessments. The integration of Large Language Models, such as ChatGPT3.5, is explored, shedding light on their potential in educational narrative assessment. The proposed methodology not only enhances understanding of class dynamics but also contributes a practical tool for educators, seamlessly translating raw data into visually compelling narratives. The paper concludes with insights from a pilot test, revealing student perceptions and addressing concerns around AI impact on dashboard utility and fairness. This research advances the intersection of data storytelling and Learning Analytics Dashboards, offering valuable insights into collaborative learning dynamics.

Keywords

Dashboards, Narrative Storytelling, Artificial Intelligence, GenAI

1. Introduction

The convergence of data storytelling and narratives within Learning Analytics Dashboards (LADs) has gained significant attention due to its potential to convey insights to non-expert audiences [1]. The use of data storytelling and narratives in education provides effective communication, personalized learning experiences, reflective opportunities, informed decision-making, and performance improvement for both students and educators. Researchers and educational stakeholders often collaborate using human-centered design approaches to align pedagogical intentions with data insights [2, 3]. While the importance of these narratives is

Joint Proceedings of LAK 2024 Workshops, co-located with the 14th International Conference on Learning Analytics and Knowledge (LAK 2024), Kyoto, Japan, March 18-22, 2024.

*Corresponding author.
adriano.pinargote@cti.espol.edu.ec (A. Pinargote); eddy.calderon@cti.espol.edu.ec (E. Calderón); kevin.cevallos@cti.espol.edu.ec (K. Cevallos); gladys.carrillo@cti.espol.edu.ec (G. Carrillo); kchilui@espol.edu.ec (K. Chiluiza); vanessa.echeverria@monash.edu (V. Echeverria)
0009–0001–8123–2684 (A. Pinargote); 0009-0004-7417-3294 (E. Calderón); 0009-0003-8581-6200 (K. Cevallos); 0000-0002-9142-6482 (G. Carrillo); 0000-0001-5992-6236 (K. Chiluiza); 0000–0002–2022–9588 (V. Echeverria)

CEUR Workshop Proceedings (CEUR-WS.org)

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

recognized in past research [4, 5], the manual generation of insights remains a time-consuming challenge. Inspired by recent exploits of Generative Artificial Intelligence (GenAI) in natural language processing [6], this paper introduces an approach for automatically generating data narratives applied in educational contexts [7]. Our approach includes segmenting behavioural data, refining GPT-3.5 comprehension prompts through prompt engineering, validating GenAI outputs, and generating a dashboard. We illustrate our approach in a pilot study with participants (students) to initially explore its usefulness, understanding and fairness in representing the data.

2. Related Works and Research Gaps

2.1. Learning Analytics Dashboards and Narratives

LADs have been widely used over the past ten years [8] as they provide a comprehensive visual overview of teacher and student performance [9] aiming at closing the feedback loop [10] and support timely interventions. Learning Analytics Dashboards (LADs) have the potential to communicate crucial insights to educators, administrators, and students, fostering informed decision-making [8]. Integrating diverse metrics, LADs provide a snapshot of academic progress, identifying patterns and trends that impact student outcomes. From tracking classroom engagement to assessing task performance, LADs are essential for targeted interventions, aiming to enhance education quality and offer personalized learning experiences [9]. However, limitations persist, including challenges in interpreting complex data and visualizations by teachers and students [11] and a gap in aligning pedagogical needs with the data [12].

In recent years, Learning Analytics (LA) researchers and designers have embraced InfoVis and visualization design principles for communicating insights to a non-expert audience [13]. Data storytelling (DS) is the art of conveying insights and information through a narrative-driven presentation of data. Several works incorporate data storytelling principles into LAD design, aiming to communicate key insights to teachers and students. This approach supports the interpretation of critical data insights, facilitating reflection and behavior adaptation for future practices [10]. Prior works have established a structured process of converting conventional visualizations into visualisations with data storytelling elements [14, 5, 15]. These dashboards aim not only to present extensive data but also to make information meaningful by delivering "one story at a time" [16] and utilizing narratives for user feedback [17]. This approach simplifies information and highlights users' key points in visualizations.

A notable gap in current research focuses on feedback mechanisms and narrative elements within LADs. While prior studies have acknowledged the potential impact of DS elements (i.e., narratives) in LADs for teaching and learning practices [14, 5, 15], there is a lack of approaches that generate these DS elements, especially if the aim is to scale the generation of such dashboards.

2.2. Generative AI (GenAI) in Education

The fundamental concept behind Generative AI lies in training models on extensive datasets, empowering them to generate original content that closely mimics human language patterns[18].

In the realm of education, there remains limited knowledge on how to optimize researchers' collaborative experiences with Generative AI.

According to [19], educational stakeholders can enhance their LADs by performing data processing and generation tasks through LLMs. Nevertheless, there exist some challenges and opportunities, such as privacy, fairness, ethics, accessibility concerns, and the inherent challenge of how to customize the LADs supported by GenAI to be more "learner" oriented. Another potential use of GenAI is in thematic analysis, where such systems offer a heightened level of user autonomy through natural language (NL) interaction, which differs from previous machine-learning and pattern-based systems [6] and the inclusion of NL processing by comparing LLMs such as GPT3 and ChatGPT to present reliable and accurate information to any audience [20] and the automatization of creating images, charts, and even maps [21] with a few NL instructions. A systematic review [22] highlighted ChatGPT's potential in healthcare research and education for performing thematic analysis tasks but emphasized the essential requirement for robust guidelines to address potential misuse and also a robust validation of the resulting coding data.

These two examples show the potential to empower researchers using GenAI for the automatic identification of behaviors. However, to our knowledge, there is a lack of approaches addressing the challenge of identifying behaviors from learning data using GenAI. Particularly, researchers should consider the alignment of pedagogical intentions (i.e., expected or salient behaviors in learning environments) with the data. In addition, given the capabilities of GenAI to extract insights and patterns from large amounts of data [19], it is valuable for LA researchers to recognize the feasibility of GenAI to automatically generate narratives of behavioural data. These GenAI narratives can uncover patterns that may be hidden due to the large amounts of data humans need to analyze and synthesize.

3. An Approach to Generating Data Narratives using GenAI

This section outlines our approach to data narrative generation within a LAD. Our goal is to translate the pedagogical intentions of the learning activity into narrative elements, emphasizing specific data aspects and providing personalized recommendations to students. The approach (Figure 1) comprises the following components:

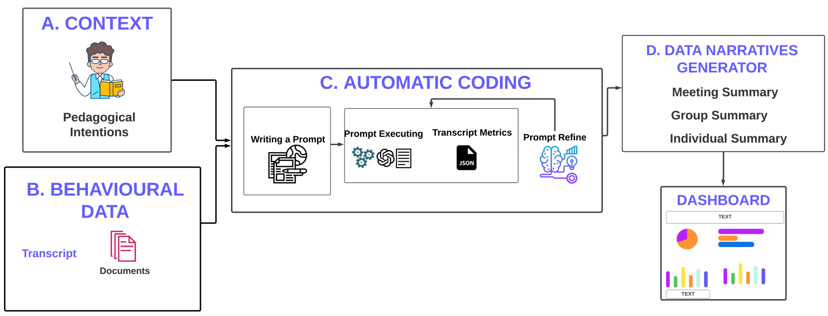

Figure 1: The proposed approach to automatically translate pedagogical intentions into data narratives.

A) Context: It is essential to understand the context of the learning environment. In our approach, this context is represented as the key pedagogical intentions that will be rendered in the dashboard. These pedagogical intentions can be derived from theory (e.g., [23]), through inquiry methods (e.g., [14]) or by using more structured approaches (e.g., [1]). These pedagogical intentions can be represented as codes or *metrics* to which we wish to draw's attention in the dashboard.

B) Behavioral Data: Our focus is to analyze behavioural data gathered from educational contexts. Here, we consider any form of data that captures human-human or human-machine interaction and that is collected through audio, video or any other device. Ideally, the goal of researchers is to code this data to reveal interaction patterns that could help understand the key pedagogical intentions in the learning context. This data could be represented as a *transcript* containing the timestamps and the content of such interactions.

C) Automatic Coding: Our aim is to feed the LLM using the pedagogical intentions (metrics) and the behavioural data (transcript) to automatically extract patterns that could be then translated into insights in the dashboard. We implemented a prompt, following the guidelines outlined in the OpenAI manual [24]. Adopting these techniques involves crafting an initial prompt, executing it using the OpenAI API, and subsequently reviewing the output to determine its correctness. This revision is an iterative process until a human (i.e., an observer or researcher) considers the output consistent and contains the desired information.

In this component, we leverage GPT 3.5 to ***automatically code the data using the key pedagogical intentions*** (metrics). In a prompt, we specify the i) topic, ii) data, iii) instruction, and iv) outcome. In the ***topic***, we describe an overview of the context, including the goal and background of the data and the pedagogical intentions (e.g., *"These are the collaborative aspects that participants should demonstrate during a collaborative activity."*). The ***data*** is the transcript split into smaller chunks to adjust to the technical specifications of GPT 3.5 (the limited use of tokens). In the ***instruction***, we describe the analysis we need for this task, which is content coding (e.g., *"Consider the data from X student. Code each utterance by identifying if the utterance contains coordination aspects."*). Finally, in the ***outcome***, we specify the desired output format (e.g., JSON) that can be used by other applications (e.g., Vue.js) or in further analysis.

D) Data Narratives Generator: In this component, we leverage GPT 3.5 to ***generate insights from the data***. In a prompt, we specify the i) topic, ii) instruction, and iii) outcome. In the ***topic***, similar to the previous component, we describe an overview of the context. We also describe an overview of the data (i.e., type, source, summary, transformations, etc) we have previously performed (as in C). In the ***instruction***, we describe the analysis we need to generate the narratives (e.g., *"Write a summary of the behaviors each participant adopted during a collaborative activity."*). Finally, similar to the previous component, we specify the desired ***outcome*** (e.g. JSON file with the narratives) that other applications will use.

Next, we illustrate our approach in the context of online collaborative learning.

4. Methods

4.1. Learning Activity and Collaboration Aspects

In the Object Oriented Programming subject, students often need to participate in collaborative activities. Students are allocated in groups (3 students per group) to solve an activity. Each group selects one of three proposed applications, which have already been curated by the teacher. Then, together, as a group, students should generate a final document specifying the objects needed to deploy the application, a description of the objects and a summary describing the most challenging object to be implemented.

From prior works, we developed six metrics [25, 23] to evaluate students' collaboration aspects exhibited during the activity, namely communication, mutual support, coordinating, work environment, commitment, and problem management. In addition, we also included on-topic/off-topic tasks to understand if students are more focused when working in the activity.

Table 1 presents the description of these metrics. These metrics and descriptions will be used as the pedagogical intentions, as described in our approach (Figure 1 - A)

4.2. Data Collection and Processing

Three researchers participated in the collaborative activity, as described above. They were asked to initiate a Teams meeting, which was video and audio recorded using an automated script. The meeting transcript, generated through Wishper [1], included timestamps and text for each participant's utterance, treated as the unit of analysis. A total of 494 utterances were identified for analysis (Figure 1 - B).

Following our approach, we automatically coded (Figure 1 - C). Meeting transcription snippets from Teams were sent to GPT 3.5. They were segmented into snippets of 900 words or less (approximately 2000 tokens) and sent to GPT 3.5. Once the entire transcript was processed, the prompt call categorized each utterance into collaboration metrics and on-topic/off-topic categories.

4.3. Coding Utterances into Collaboration Metrics

Each utterance was coded into one or several metrics, meaning that multiple collaborative aspects can appear within a single utterance. In addition, each utterance was also categorized into on-topic/off-topic. We performed a human validation with three raters and employed Krippendorff's Alpha and Cohen's Kappa to assess the inter-rater reliability. This validation was carried out in two parts:

- **Human Coding:**
  Three coders independently coded each utterance. The inter-rater reliability analysis required a collective commitment of 165 hours from three human observers —50, 55, and 60 hours each— to categorize utterances according to the specified metrics. At the end of this task, the coders participated in a discussion session to reach a consensus on the final code per each utterance. A threshold of 0.61 indicates a substantial agreement between

[1] https://openai.com/research/whisper

Table 1
Inter-rater reliability Results

Metric	Description	3 coders (Krippendorff)	Human Vs GenIA (Kappa)
On Topic	Comments relevant to the main topic	**0.72**	**0.661**
Off Topic	Comments unrelated to the main topic	**0.681**	**0.680**
Communication	Effectiveness of verbal communication among team members	0.507	0.557
Mutual Support	Assisting when facing challenges	**0.734**	**0.749**
Coordination	Team's ability to work together efficiently	**0.802**	**0.872**
Work Environment	Creates a conducive and inclusive work environment within the team.	0.443	**0.790**
Commitment	Measures team members' dedication and engagement towards shared goals	0.503	**0.728**
Problem Management	Effectively manage disagreements and problems during meeting	0.558	**0.668**

raters. As listed in Table 1, coders achieved a substantial agreement in on-topic/off-topic coordination and mutual support. In contrast, a moderate agreement was achieved in the rest of the metrics.

- **GenAI:** The previously mentioned prompt call was employed for the automatic coding of each utterance. Notably, the GenAI accomplished this task within 5 minutes. We computed the agreement between human coding and GenAI results, utilizing Cohen's Kappa. Remarkably, the GenAI demonstrated substantial agreement in on-topic/off-topic categories and all collaboration metrics, except for Communication, where moderate agreement was observed (Kappa=0.557).

4.4. Data Narratives

From the data resulting in the automatic coding, we calculated the corresponding percentages for each collaboration aspect per group and per individual. This information was used in the data narratives generator (Figure 1 - D). We generated 1) ***a summary of the activity***, describing the main insight of the meeting and the role of each student during the activity; 2) ***a group summary***, describing the overall group's behaviors concerning the collaboration metrics; and 3) ***individual feedback,*** describing the main insight and areas of improvement per each student according to the collaboration metrics.

4.5. Dashboard Prototype

Based on the outputs from the automatic coding and data narratives (Figure 1 - C & D), a high-fidelity dashboard prototype was developed using Vue.js. The web application receives two JSON files to generate the charts. The dashboard has three main elements, as illustrated in Figure 2.

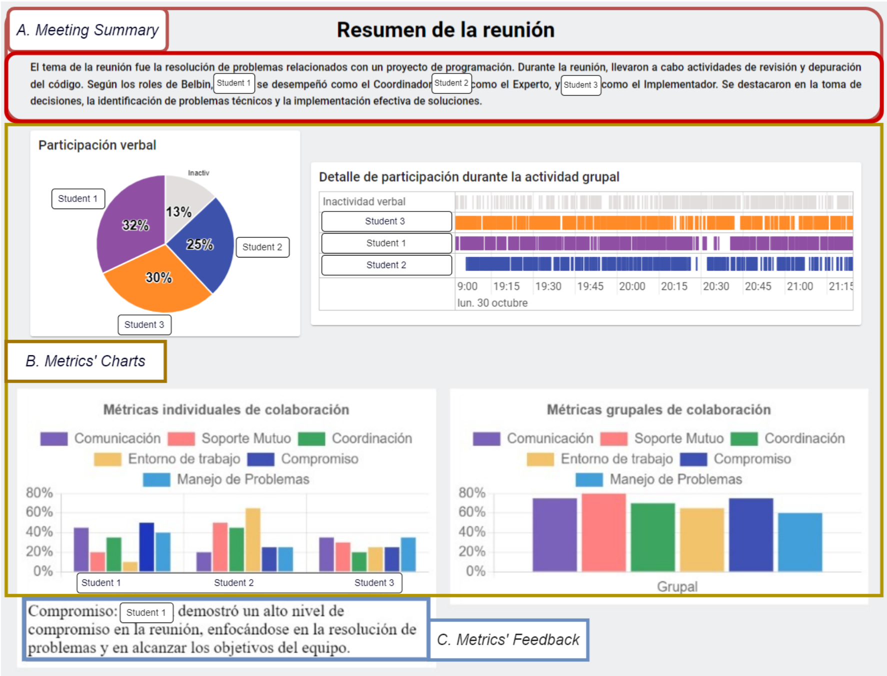

Figure 2: GenAI-powered dashboard: A) Meeting summary, B) Metrics' Charts and C) Metrics' Feedback

The dashboard displays graphs and text detailing participants' assessments for each metric, using distinctive colors for clarity. The primary goal is to provide an overview of the meeting, identifying prominent roles and areas for improvement. It starts with a general summary (Figure 2A), followed by detailed individual and group metrics (Figure 2B). This approach offers a comprehensive view, enabling students and teachers to identify areas for improvement and aspects needing consolidation and strengthening (Figure 2C) based on the six metrics.

5. Pilot Study with Students

5.1. Participants and Procedure

A pilot study with third-year bachelor students from computer science and mechatronics undergraduate programs was conducted to explore the usefulness, understanding, fairness and perceived impact of the dashboard. 19 students (3 female) participated in this pilot study. Students were exposed to the dashboard, similar to the example in Figure 2. The task consisted of 1) exploring each part of the dashboard (A, B and C) using a think-aloud protocol and 2) answering Likert-scale (1-5; 1 being the lowest value and 5 being the highest value), yes/no and open-ended questions based on their perceptions. Six questions (Q# represents the question' number) were focused on gathering perceptions (usefulness, understanding, fairness) from the LAD. The students did not know the LAD was generated using GenAI. After each Likert scale and yes/no questions, students were asked to elaborate on their responses.

Table 2
Interview Questions and Responses (19 participants)

Q#	LAD Section	Type	Mean	Median	Min	Max	Std Dev
1	Metrics' Charts	Usefulness	4.42	4.00	3.00	5.00	0.61
2	Metrics' Feedback	Usefulness	4.32	5.00	3.00	5.00	0.82
3	Summary	Understanding	4.32	5.00	3.00	5.00	0.82
4	Summary	Usefulness	4.32	5.00	1.00	5.00	1.00
#	Yes NO Question		YES	NO			
5	Metrics' Charts & Feedback	Fairness	13	6			
6	Summary	Fairness	12	7			

5.2. Preliminary Analysis and Results

For the Likert scale questions, we calculated basic statistics (mean, median, min, max, std. dev.). For the open-ended questions, we searched for quotes that could help us understand the values and challenges when exploring the dashboard. Table 2 summarizes the findings of the pilot study. Students' responses (S# represents students' comments) were gathered according to their experience with the LAD:

- **Metrics' Charts and Feedback:** In terms of usefulness, students indicated a high perception on these chart and narratives in the metrics chart (Q1) (Figure 2-B). Students appreciated the insights gained from charts, suggesting a valuable tool for enhancing their understanding of collaborative aspects. S10 mentioned: *"The metrics of what each participant truly contributed are valuable to me and how I can improve in the following meetings."*. Regarding the metrics' feedback (Q2), it has a wider standard deviation (0.82), this means that students found it useful but some may hold a diverse perspective on the usefulness of this feedback (Figure 2C). S2 said: *"I believe they (metrics' feedback) are redundant because if I already have a percentage (metric chart), the textual summary can be inferred."* In fairness (Q5), most students (65%) perceived that the metrics' feedback are fair, because they recognize each student contribution, S16 says: *"They (metrics' feedback) not states non-participation but ensures to present the qualities and both strengths and weaknesses."*, but S12 (and 35%) opposes: *"I don't perceive them (metrics' feedback) as directly related to the collaboration metrics percentages (Figure 2B). It would even be unfair to compare students in this context."*
- **Summary:** Students generally found the summary (Figure 2A) easy to understand (Q3) because they were able to identify their roles. They also valued that the summary highlighted what had been discussed at the meeting. S5 expressed: *"Because it can provide self-feedback, I can reflect on the roles I play in a team. I realized that sometimes I assume a particular role frequently."*. However, in terms of its usefulness (Q4), while most students agreed that this summary and it's narrative is useful, one student perceived the opposite – affecting the standard deviation. This student explained that: *"If the (summary) feedback is provided to us by the teacher (in-person), it has more significance."*. In terms of fairness (Q6), most of the students (65%) agreed that the general summary fairly describes the

participation of all students. As expressed by S16: *"Capturing the entire context of the activity is crucial. It doesn't imply a lack of participation; rather, it ensures a comprehensive presentation of qualities, strengths, and weaknesses."*. In contrast, 35% expressed that the information provided is not fair. S12 remarked: *"Assigning roles (to students) might seem unfair, someone could feel uncomfortable being consistently confined into a specific role."*

6. Discussion and Future Work

In the challenge of moving from raw data to an interpretation of results, the approach followed in this paper resembles prior works [26], where data from meeting recordings were utilized to create a dashboard employing social network analysis, providing team members with insights into their communication behavior. Our approach uses the same idea but differs by automating the process of generating insights using GenAI [19]. Through our approach, we effectively analyzed the quality of collaboration by measuring collaborative aspects from the transcript, reducing time and use of human resources. Preliminary responses from students show promising results about the potential value of these automatic process, but also highlight the challenges of using GenAI, in particular, when discussing the fairness of representation. Researchers are invited to investigate ethics, reliability, trustworthiness, and safe principles inherent to intelligent LAD systems [27].

One challenge of this research was to handle a substantial volume of data extracted from human-human and human-computer interactions, in our case, through speech transcripts. One transcript can contain 50 pages of text, which is a large amount of data for an LLM (i.e., GPT 3.5). Effectively processing this large dataset presented difficulty in achieving accurate results, given that the more data, the more difficult it becomes for model to understand the whole context. Our solution uses various Prompt Engineering techniques [24] and sheds light on the automatic generation of narratives that communicate insights, a challenge highlighted in prior LAD literature[8]. However, this approach opens up new research venues and opportunities to address issues such as prompt and data quality validation, inviting researchers to investigate and develop human-AI collaboration approaches to overcome these challenges [18].

AI denotes remarkable efficiency in recognizing patterns [18], completing assessments in minutes compared to humans, highlighting its potential for streamlining evaluations, but the lack of transparency in the "blackbox" process is remarkable as it is a closed-source system; the inappropriate use of LLMs could raise legitimate concerns about data privacy [28], which brings us to the issue of security [29], security risks in ChatGPT commonly include command injection, data poisoning, privacy leaks, malicious content generation and ethical risks related to bias, highlighting the need for robust security and ethical measures in its development and use, bearing in mind that this management [30] of the information is extremely isolated from the user.

Choosing Cohen's Kappa and Krippendorf for this study allowed us to highlight the ability to measure inter-rater agreement robustly [31], the substantial agreement among human observers and moderate to substantial agreement with AI assessments suggest reliability convergence, challenging conventions and emphasizing the need for collaborative human-AI evaluation in education [6, 18]; the inclusion of AI would be sooner a "normal" path to be taken to potentiate the education field; however, there is the need of validation until a reliable standard is reached.

7. Conclusion

This article introduces an automated approach employing GenAI for the creation of data narratives in the context of collaborative learning through the use of LADs. The research showcases the potential of AI in streamlining evaluation processes and underscores the significance of a collaborative synergy between human and AI evaluators. Additionally, it highlights the crucial role of exploring and utilizing various tools within the LLM environment for achieving precise outcomes. The overall implications of the findings suggest that the integration of data storytelling and narratives into LADs holds promise for educational stakeholders and non-expert audiences, offering valuable insights and improving comprehension of collaborative endeavors.

References

[1] G. M. Fernandez-Nieto, R. Martinez-Maldonado, V. Echeverria, K. Kitto, D. Gašević, S. Buckingham Shum, Data storytelling editor: A teacher-centred tool for customising learning analytics dashboard narratives, in: LAK24: 14th International Learning Analytics and Knowledge Conference, 2024.

[2] D. Wang, H. Han, Applying learning analytics dashboards based on process-oriented feedback to improve students' learning effectiveness, British Journal of Educational Technology 51 (2020) 555–569.

[3] V. Echeverria, R. Martinez-Maldonado, R. Granda, K. Chiluiza, C. Conati, S. Buckingham Shum, Driving data storytelling from learning design, in: Proceedings of the 8th International Conference on Learning Analytics and Knowledge, LAK '18, Association for Computing Machinery, New York, NY, USA, 2018, p. 131–140.

[4] B. Dykes, Data storytelling: What it is and how it can be used to effectively communicate analysis results, 2015.

[5] R. Martinez-Maldonado, V. Echeverria, G. Fernandez, S. Buckingham Shum, From data to insights: A layered storytelling approach for multimodal learning analytics, 2020, pp. 1–15. doi:10.1145/3313831.3376148.

[6] J. V. Pavlik, Collaborating with chatgpt: Considering the implications of generative artificial intelligence for journalism and media education, Journalism & Mass Communication Educator 78 (2023) 84–93.

[7] C. K. Y. Chan, W. Hu, Students' voices on generative ai: perceptions, benefits, and challenges in higher education, International Journal of Educational Technology in Higher Education 20 (2023) 43.

[8] I. Jivet, M. Scheffel, H. Drachsler, M. Specht, Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice, in: É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, M. Pérez-Sanagustín (Eds.), Data Driven Approaches in Digital Education, Springer International Publishing, Cham, 2017, pp. 82–96.

[9] W. Matcha, N. A. Uzir, D. Gašević, A. Pardo, A systematic review of empirical studies on

learning analytics dashboards: A self-regulated learning perspective, IEEE Transactions on Learning Technologies 13 (2020) 226–245.
[10] W. Greller, H. Drachsler, Translating learning into numbers: A generic framework for learning analytics, Educational Technology Society 15 (2012) 42–57.
[11] L. Corrin, Evaluating Students' Interpretation of Feedback in Interactive Dashboards, 2018, pp. 145–159.
[12] K. Verbert, X. Ochoa, R. D. Croon, R. A. Dourado, T. D. Laet, Learning analytics dashboards: the past, the present and the future, Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (2020).
[13] C. N. Knaflic, Storytelling with data: A data visualization guide for business professionals (12 ed.), Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], John Wiley Sons, New Jersey, USA, 2017.
[14] V. Echeverria, R. Martinez-Maldonado, S. Buckingham Shum, K. Chiluiza, R. Granda, C. Conati, Exploratory versus explanatory visual learning analytics: Driving teachers' attention through educational data storytelling, Journal of Learning Analytics 5 (2018).
[15] S. Pozdniakov, R. Martinez-Maldonado, Y.-S. Tsai, V. Echeverria, N. Srivastava, D. Gasevic, How do teachers use dashboards enhanced with data storytelling elements according to their data visualisation literacy skills?, in: LAK23: 13th International Learning Analytics and Knowledge Conference, LAK2023, New York, NY, USA, 2023, p. 89–99.
[16] G. Fernandez-Nieto, V. Echeverria, S. Buckingham Shum, K. Mangaroska, K. Kitto, E. Palominos, C. Axisa, R. Martinez-Maldonado, Storytelling with learner data: Guiding student reflection on multimodal team data, IEEE Transactions on Learning Technologies PP (2021) 1–1.
[17] G. Fernandez Nieto, K. Kitto, S. Buckingham Shum, R. Martinez-Maldonado, Beyond the learning analytics dashboard: alternative ways to communicate student data insights combining visualisation, narrative and storytelling, in: LAK 2022 Conference Proceedings, USA, 2022, pp. 219–229.
[18] L. Yan, V. Echeverria, G. F. Nieto, Y. Jin, Z. Swiecki, L. Zhao, D. Gašević, R. Martinez-Maldonado, Human-ai collaboration in thematic analysis using chatgpt: A user study and design recommendations, 2023.
[19] L. Yan, R. Martinez-Maldonado, D. Gasevic, Generative artificial intelligence in learning analytics: Contextualising opportunities and challenges through the learning analytics cycle, 2023.
[20] P. Maddigan, T. Susnjak, Chat2vis: Generating data visualisations via natural language using chatgpt, codex and gpt-3 large language models, 2023. doi:10.48550/arXiv.2302.02094.
[21] R. Tao, J. Xu, Mapping with chatgpt, ISPRS International Journal of Geo-Information 12 (2023). URL: https://www.mdpi.com/2220-9964/12/7/284. doi:10.3390/ijgi12070284.
[22] M. Sallam, Chatgpt utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns, Healthcare 11 (2023).
[23] A. Meier, H. Spada, N. Rummel, A rating scheme for assessing the quality of computer-supported collaboration processes, International Journal of Computer-Supported Collaborative Learning 2 (2007) 63–86.
[24] OpenAI, Openai, six strategies for getting better results, https://platform.openai.com/docs/

guides/prompt-engineering/six-strategies-for-getting-better-results, 2023.
[25] V. Echeverria, P. Garaizar, J. Garcia-Zubia, Dimensions-an exploratory evaluation of a collaboration feedback report, in: Proceedings of the 12th International Conference on Learning Analytics and Knowledge (LAK22), 2022, pp. 480–484.
[26] T. Spielhofer, R. Motschnig, Developing teams by visualizing their communication structures in online meetings, Multimodal Technologies and Interaction 7 (2023).
[27] R. Alfredo, V. Echeverria, Y. Jin, Z. Swiecki, D. Gasevic, R. Martinez-Maldonado, Slade: A method for designing human-centred learning analytics systems, 2024.
[28] M. Gupta, C. Akiri, K. Aryal, E. Parker, L. Praharaj, From chatgpt to threatgpt: Impact of generative ai in cybersecurity and privacy, IEEE Access 11 (2023) 80218–80245. doi:10.1109/ACCESS.2023.3300381.
[29] X. Wu, R. Duan, J. Ni, Unveiling security, privacy, and ethical concerns of chatgpt, Journal of Information and Intelligence (2023). URL: https://www.sciencedirect.com/science/article/pii/S2949715923000707. doi:https://doi.org/10.1016/j.jiixd.2023.10.007.
[30] L. Ayinde, M. P. Wibowo, B. Ravuri, F. B. Emdad, Chatgpt as an important tool in organizational management: A review of the literature, Business Information Review 40 (2023) 137–149. URL: https://doi.org/10.1177/02663821231187991. doi:10.1177/02663821231187991. arXiv:https://doi.org/10.1177/02663821231187991.
[31] Z. Xiao, X. Yuan, Q. V. Liao, R. Abdelghani, P.-Y. Oudeyer, Supporting qualitative analysis with large language models: Combining codebook with gpt-3 for deductive coding, in: 28th International Conference on Intelligent User Interfaces, ACM, 2023, pp. 1–4.