Click to Empower

Literature Review on (LMS-based) Learning Analytics for Student Well-Being

Report 1 - June 2024

Drs. Ran Zhang, Dr. Uwe Matzat, Dr. Rianne Conijn, and Prof. dr. Chris Snijders

EINDHOVEN UNIVERSITY OF TECHNOLOGY

4TU. CENTRE FOR ENGINEERING EDUCATION

Introduction

Student well-being deterioration observed these years has raised concern (Downs et al., 2017; Pedrelli et al., 2015), as they correlate with poor academic performance, compromised physical health, and increased substance use problems (Pascoe et al., 2020). Two main approaches to well-being exist: *hedonic well-being*, which emphasizes positive emotions and the absence of negative emotions, and *eudaimonic well-being*, which focuses on perceiving meaning and functioning in life (Disabato et al., 2016; Hossain et al., 2023). In the educational context, researchers typically use hedonic, eudaimonic, or an integrated approach to define student well-being (Hossain et al., 2023). A hedonic approach considers students' positive and negative emotions, such as satisfaction, enjoyment, and the absence of worries, anxiety, and stress. A eudaimonic approach concentrates on students' self-assessments of academic functioning, such as engagement, self-efficacy, time management, and social connectedness. An integrated approach then combines elements from both. In this report, we adopt an integrative approach, combining hedonic (e.g., anxiety, stress, depression) and eudaimonic (behavioral/functional, e.g., time management, procrastination, resilience) aspects of well-being. Specifically, we identify approaches towards measuring student well-being.

There is a growing interest in utilizing clickstream data from Learning Management Systems (LMS) to gain insights into students' learning processes and provide better support, making it an important area in Learning Analytics research (Hellas et al., 2018). While LMS data have been successful in predicting academic performance, its potential for monitoring student well-being is still under exploration. In this report, we hence aim to explore: What has been accomplished in this much-needed field thus far?

Research Questions

To better understand the state of the art in using LMS data to identify and predict student well-being and to better identify other potential research directions and challenges in this field, a literature review was conducted to answer the following questions:

1. What is the purpose of these predictive models or the construction of LMS indicators for student well-being?
2. What student well-being constructs are targeted in these models?
3. What clickstream indicators are used, and which ones appear to be effective?
4. What predictive modeling methods are used, and what is the performance? What can be concluded looking at the performance?
5. What are the challenges of using LMS-based learning analytics to monitor student well-being?

Method

A comprehensive search was conducted in January-May 2024 in IEEE Xplore, Web of Science, Scopus, and PubMed using the following search query:

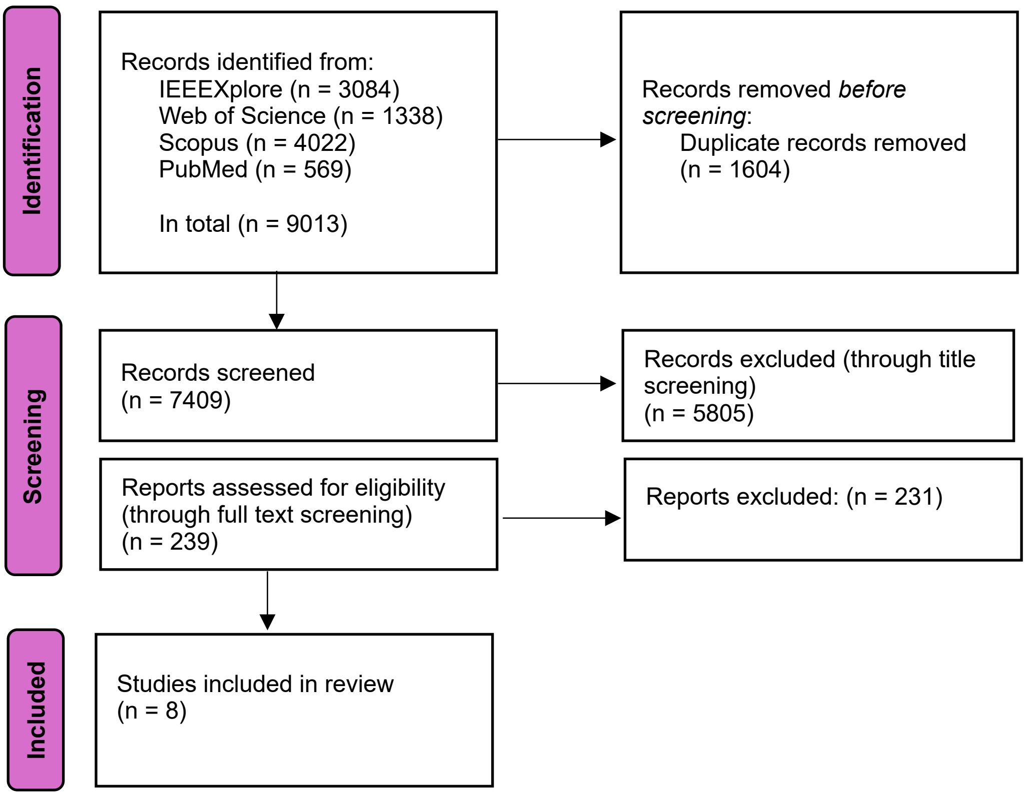
TS* = (“learning analytics” or “education* data mining” or “learning management system” or “LMS” or “Canvas” or "Blackboard" or "Moodle" or "e-learning"). and TS=(“depression” or "stress" or "resilience" or “mental health” or "well-being"). [*TS= topic, including title, abstract, keyword.]

The inclusion criteria for the studies were as follows: (1) The study must utilize information from an online learning management platform, such as MOOC, Canvas, Moodle, or similar systems. (2) The prediction target or the goal of the LMS indicator construction should be (at least partly) to better understand/predict student well-being, rather than solely on performance or dropout. (3) The full text must be available. (4) The full text must be in English, Dutch, German, or Chinese to ensure comprehension by the report authors.

For the review process, one reviewer conducted the screening and data selection manually without automation tools. The variables collected, based on the research questions, included the purpose of the study, targeted student well-being constructs, LMS indicators constructed, predictive techniques employed, and challenges encountered. Figure 1 provides an overview of the screening process. Initially, 9,013 items were identified using the search query. After removing duplicates, 7,409 items remained. A title screening was then conducted, reducing the number to 239 items. Following abstract and full-text screening of these 239 items, only eight papers remained.

Figure 1

Screening process



Results

Eligible papers

In total, eight papers were deemed relevant for the current literature review. Within the eight papers, two categories of related papers were identified. In the first category, LMS information was used mainly to understand and measure certain mental well-being constructs, but no prediction was made. For example, researchers might use an online login time pattern to infer sleeping problems. While in the second category, LMS information was not only used to construct well-being indicators, but also predictive models were fitted and evaluated. For example, in this case, sleeping problems would be measured through self-report as the true label, and then LMS information would be used to predict the self-reported sleeping problems through predictive modeling. There are four papers in each category.

Category 1: Measuring/representing well-being through LMS data

In the first paper, Cervera-Mercadillo and Lara (2020) proposed a LMS clickstream analysis scheme aimed at capturing student procrastination. They used Moodle as the example platform. For each student and each task, they constructed a "degree of delay in task completion" indicator, coded as follows: negative for early completion (eager), positive for late completion (delay), and zero for on-time completion (on time), according to the instructor's recommended completion time. The data was aggregated weekly by each task category (e.g., viewing presentations, submitting assignments). With this aggregated information, the authors claimed that this analysis framework could help better predict academic performance and dropout rates, although in this paper it has not been tested whether it improved the prediction results.

In the second paper, Abdullah (2021) explored the potential of LMS data to identify possible sleep disruptions/sleep disorders. By examining the time patterns (whether it is outside usual working hours) of login and course page viewing behavior of students and instructors, Abdullah argued that online teaching and learning could potentially reveal patterns indicative of sleep disorders.

In the third paper, Rodrigues and colleagues (2013) estimated student stress during Moodle-based learning sessions using keypress and mouse movement data. They compared two task conditions: one with a time limit and higher stakes (telling participants that task performance would significantly impact student grades), inducing more stress, and another without a time limit and lower stakes. The data collected included the number of keypresses and mouse movements, aggregated for each task completion session. The results show that stressed students exhibited more mouse usage and keyboard pressing, greater hesitation and jitteriness in mouse clicks and scrolls, and more intensive use of the backspace key, supporting the possibility of integrating mouse and keystroke data from LMS based task environment into real-time stress detection.

In the fourth paper, Yao and colleagues (2020) proposed a sophisticated temporal model to capture student procrastination. They used temporal patterns rather than static measures, and they addressed the interdependence between the time sequences of different activities (e.g., past and future activities from different modules) by categorizing online learning activities into lecture-related (L), assignment-related (A), and discussion-related (D) activities. The Hawkes process model along with some other temporal models were used to fit the data, capturing the density and burstiness of student activities. According to the results, the multidimensional Hawkes model fitted the time sequence data the best, and the parameters of the fitted Hawkes model could represent entities such as intensity, frequency, regularity of the study activity. Moreover, a negative correlation between delay parameter and student grades was observed, further supporting its criterion validity in improving LMS-based academic performance prediction.

Category 2: Predicting well-being through LMS data

In the fifth paper, Maaliw and colleagues (2022) developed a machine learning framework to identify student grit through online learning behavior. They mapped Moodle indicators to sub-components of the grit construct in a one-on-one manner: self-motivation (indicated by number of quiz attempts), self-regulation (hours spent online outside the course schedule), diligence (submission time), perseverance (number of attempts to complete a challenging coding task), and sustained interest (consistency of logins, modeled through dynamic time warping technique). Then, both cluster analysis and classification analysis were used to predict grit: Firstly, these five indicators were clustered separately into high/low grit clusters, and according to these five clusters, a majority voting system determined the final high/low grit classification for each student. Later, a series of classification algorithms were used to predict grit through these indicators and achieved accuracy between 0.812 and 0.921, with random forest performing the best. Clustering and classification results showed high consistency with each other, as well as with grit questionnaire results (90.55% to 94.07%), proving the success of both techniques. From this paper, it seems that explicitly mapping students' online study behaviors to the structure of the latent well-being construct can be a good idea when predicting complex constructs like grit. Moreover, it shows how unsupervised and supervised learning approaches can be combined in predicting well-being through LMS data.

In the sixth paper, Guo and colleagues (2022) predicted student mental health (measured by SCL-90) using information from Chaoxing, an LMS used in China. They included final course grades, student photos, and social relations in their predictive model, employing a three-layer deep neural network, which achieved an accuracy of 0.845. This study demonstrated the possibility of fusing other student life information with LMS data for predicting student well-being; however, the information from LMS in this study is rather limited, as only grade was incorporated.

In the seventh paper, Aljarallah and colleagues (2022) attempted to predict depression, anxiety, and stress using LMS data. However, details about the measurement of these constructs and the specific LMS data used were limited, as the online open access dataset they specified in the paper could not be found at the time of writing this report. Despite this, the reported classification accuracy ranged from 0.798 to 0.991 across different algorithms.

In the eighth paper, Zhao and colleagues (2023) aimed to predict student procrastination using the Chinese LMS Chaoxing. They constructed three indicators: inactive time, active time, and idle time across three categories of online activities: course content, course discussions, and assignments. These indicators were calculated based on content publish time, initial viewing time, completion time, and task deadlines, and then aggregated for each individual. A backpropagation neural network was used for predictive modeling, along with other classical classification algorithms, with accuracy ranging from 0.785 to 0.937 (the highest achieved by the BP neural network). This paper showcases the potential in LMS based well-being prediction when LMS clickstreams indicators were thoughtfully constructed and then combined with more advanced machine learning algorithms.

Research Questions

RQ1. Purpose of these studies

The purposes of these reviewed studies are as follows: To understand and monitor student well-being with the goal of better predicting academic performance and dropout (Cervera-Mercadillo & Lara, 2020; Yao et al., 2020); To monitor student well-being with the aim of offering interventions to help students build non-cognitive abilities (Maaliw et al., 2022; Rodrigues et al., 2013; Zhao et al., 2023); To identify possible mental health crises for timely intervention and general monitoring purposes (Aljarallah et al., 2022; Guo et al., 2022); and lastly to identify possible adverse impacts of online learning on students' well-being (Abdullah, 2021).

RQ2. Targeted student well-being constructs and their measurement

Among the eight studies reviewed, three focused on procrastination (Cervera-Mercadillo & Lara, 2020; Yao et al., 2020; Zhao et al., 2023); four focused on mental health-related well-being constructs such as depression, stress, anxiety, or general mental health problems (Abdullah, 2021; Aljarallah et al., 2022; T. Guo et al., 2022; Rodrigues et al., 2013); and one focused on grit (Maaliw et al., 2022). Among the four prediction papers, most utilized self-report questionnaires to measure the mental health constructs (Aljarallah et al., 2022; T. Guo et al., 2022; Maaliw et al., 2022), while one study did not explicitly specify its measurement method (Zhao et al., 2023). From these papers, it is apparent that constructs with a time component are selected more often (procrastination, grit, sleep behavior).

RQ3. LMS indicator construction and comparison

Students' interactions with time-bound tasks and their time patterns (whether completed before the recommended deadline or not) are frequently deemed meaningful and are extracted in most studies. However, the presentation and integration of this information into the final models vary slightly across the papers in three distinct ways: (1) Level of reservation of the original temporal pattern: This ranges from preserving the raw time sequence to aggregating the data for each student through the whole semester (e.g., total number of attempts on an assignment). (2) Representation of late assignment: For each assignment, the representation can be categorical (e.g., late, just on time, early) or more continuous (e.g., the actual time gap between submissions and deadlines). (3) Covered domain of LMS content: Some studies only look at the time pattern about task completion or assignment submission, however, some also look at the time pattern and its relative position with deadline/suggested time for for other aspects than assignments, such as course content views, and course discussions.

The consistency or regularity of students' learning behavior or login frequency is also commonly mentioned (Maaliw et al., 2022; Yao et al., 2020). In these papers, the consistency or regularity was addressed using dynamic time warp technique, or with Hawkes Process model. Additionally, grades have been included in one study (T. Guo et al., 2022), while the number of attempts on ungraded quizzes and tasks was used in another (Maaliw et al., 2022). Furthermore, the number of mouse and keyboard movements appears to increase under stress (Rodrigues et al., 2013), and LMS login times can indicate irregular sleeping hours (Abdullah, 2021).

However, it is difficult to conclude which indicators might be the most effective for well-being monitoring because these studies do not use a comprehensive selection of indicators, and only one paper explicitly discussed the relative importance of different indicators in their predictive models: using the permutation feature importance metric, Maaliw and colleagues (2022) found that in their best performing random forest model in grit predictions, perseverance (indicated by number of attempts to complete a challenging coding task) is the best predictor, followed by sustained interest (indicated by consistency of logins, modeled through dynamic time warping technique). Since they only mapped indicators according to the measurement of the well-being construct (in this case, grit), still no conclusion can be confidently made about the cross-domain importance of certain indicators.

RQ4. Predictive modeling technique, the performance, and research/practical implications

Classification techniques are predominantly used in LMS-based well-being monitoring. Some studies utilize classic classification algorithms, e.g., logistic regression, random forest, while others develop more innovative algorithms and compare their performance with classic ones. Since all four reviewed papers with predictive modeling were classification studies, metrics used to evaluate model performance include F1 score, precision, sensitivity, accuracy, confusion matrix, ROC, and AUC. An accuracy ranging from 0.785 to 0.991 (with the highest value achieved in the prediction by Aljarallah and colleagues (2022), which lacked sufficient information on LMS indicator construction) was reported across the studies. This represents rather high performance. Comparing these results to the previous prediction results from our dataset (see Corona Monitoring Report 2-C), although the performance of a classifier and a regressor cannot be directly compared (as they serve different purposes, classification into categories and prediction of exact values), the classification results within these studies appear strong relative to our regression results.

The high performance reported in some papers could be attributed to several factors (1) Reduced noise in LMS indicators. Some studies employed well-defined, nuanced, temporal, carefully mapped-out indicators, sometimes combined with dimension reduction technique, which improved the quality of the individual indicators and reduced the overall number of indicators. Moreover, advanced feature selection techniques were also used in some studies to make sure the quality of the finally selected indicators. Additionally, 2 out of the 3 studies that provided details on LMS indicator construction used records from within the same course, eliminating between-course variance and resulting in lower noise. (2) Reduced noise in target variables. Making some student well-being measures binary might after all be beneficial and worth a try if we have good reasons for the threshold setting. Firstly, the underlying distribution could be bimodal. Secondly, continuous measures of well-being might contain measurement errors and be rather noisy. In this case, by categorizing them into binary or categorical variables, the noise is reduced. (3) Advanced machine learning algorithms. The use of more sophisticated machine learning algorithms likely contributed to the improved performance. These factors likely contribute to the higher classification performance observed in the reviewed studies compared to our previous regression results and can inform the optimization of future data analysis techniques.

No definitive conclusions can yet be drawn about the conditions (pre-processing methods used; indicators used; predictive models used etc.) that affect prediction performance, given the idiosyncrasy as well as the limited number of these studies. Further research is needed to determine the factors that influence the accuracy and generalizability of these predictive models.

Nevertheless, the constructs discussed in these papers can be predicted with relatively high accuracy. Given this prediction quality demonstrated, it can be concluded that for certain student well-being constructs (e.g., procrastination, grit, mental wellness), the classification accuracy is sufficiently high for practitioners and educators to act on these prediction results, such as to screen students who may need help or are at risk of delay and intervene in a timely manner.

RQ5. The challenges of LMS-based learning analytics in student well-being monitoring

From the reviewed studies, it can be inferred that in LMS-based well-being monitoring, a significant challenge lies in extracting meaningful signals from the rich and potentially noisy LMS data while minimizing information loss. LMS data are comprehensive since it contains all time-stamped learning activities. However, this abundance of data makes it challenging to identify useful patterns and meaningful information.

First, it is delicate to accurately interpret and handle time patterns. Student activity can depend on many factors, such as the course setup or even the setup of another different course the student is taking. However, no paper in this review has yet addressed the variance of the displayed online study behavior between courses, nor the variance of the predictive power of the same indicator between different courses. Second, study patterns or regularities might not be consistent across different domains. For example, students might procrastinate on submitting assignments or engaging with assignment-related content but not on viewing course slides, which are easier and have a lower threshold. These nuances could be easily overlooked if not carefully considered. Moreover, the more detailed the data, the more challenging the analysis, making predictive modeling more prone to overfitting. However, oversimplification then again can lead to significant information loss, preventing the full utilization of the advantages of LMS data.

These challenges may explain why various advanced methods were employed in the reviewed papers. For example, to address the noisy nature of the data, some researchers employ cluster analysis with majority vote techniques. Moreover, to treat the noisy time-related patterns, instead of aggregating all LMS-logged events, some studies tried to retain the original time sequence of these events. Additionally, when treating LMS data as time sequences, different categories of activities may exhibit varied development patterns or distributions, further complicating the analysis. In such cases, some researchers separate time sequences according to each activity category and use computational modeling to address potential interactions among these categories.

Despite the reported model evaluation results, no definitive conclusion can be drawn about the optimal treatment of time-sequence data so far, since the number of reviewed papers was rather limited.

Conclusion

Currently, efforts are underway to integrate LMS data to enhance the representation and prediction of student well-being. These efforts aim to better explain variations in academic performance, better monitor and support students' non-cognitive abilities, and monitor mental health issues. Constructs of interest in well-being include academic well-being, mental health status, stress, and sleep problems. Procrastination has gained quite some attention. Key LMS indicators often used in this context include the timeliness of task completion and the regularity of study activities, which again also correlate with procrastination behavior a lot. For prediction, techniques such as cluster analysis, classic classification algorithms, and innovative classification algorithms are employed. For the challenge of LMS-enabled student well-being prediction/monitoring, it seems that the challenge lies in extracting meaningful signals from the rich and potentially noisy LMS data while minimizing information loss, especially for the rich time patterns.

There is still uncertainty regarding which indicators might be most effective in predicting well-being, as few LMS indicators have been extensively studied. In addition, no study has looked at how the between-course variance of these LMS indicators play a role in well-being prediction (as is done for academic performance, e.g., see Conijn et al., 2017). Yet these insights would help immensely when practitioners want to make sense of the LMS information to inform practical intervention or to inform course design. Furthermore, most studies focus on only one or, at most, three well-being constructs, making it difficult to determine which well-being indicators are the most feasible, therefore worth the effort the most, to predict. To conclude, future studies should keep exploring useful LMS indicators across different well-being aspects, while at the same time also trying to identify feasible and useful student well-being constructs to be predicted and keep exploring effective ways to retain useful time-related patterns from LMS to feed the predictive models.

References

Abdullah, A. (2021). Sleep Behaviour and Online Engagement in Learning Management System at Higher Education During COVID-19 Pandemic. *2021 International Conference on Software Engineering & Computer Systems and 4th International Conference on Computational Science and Information Management (ICSECS-ICOCSIM)*, 143–148. https://doi.org/10.1109/ICSECS52883.2021.00033

Aljarallah, N., Dutta, A., Alsanea, M., & Rahaman, A. (2022). Intelligent Student Mental Health Assessment Model on Learning Management System. *Computer Systems Science and Engineering*, *44*(2), 1853–1868. https://doi.org/10.32604/csse.2023.028755

Cervera-Mercadillo, F., & Lara, J. A. (2020). *A method for generating features representing the students' degree of anticipation/delay to complete assignments*.

Conijn, R., Snijders, C., Kleingeld, A., & Matzat, U. (2017). Predicting Student Performance from LMS Data: A Comparison of 17 Blended Courses Using Moodle LMS. *IEEE Transactions on Learning Technologies*, *10*(1), 17–29. https://doi.org/10.1109/TLT.2016.2616312

Disabato, D. J., Goodman, F. R., Kashdan, T. B., Short, J. L., & Jarden, A. (2016). Different types of well-being? A cross-cultural examination of hedonic and eudaimonic well-being. *Psychological Assessment*, *28*(5), 471–482. https://doi.org/10.1037/pas0000209

Downs, A., Boucher, L. A., Campbell, D. G., & Polyakov, A. (2017). Using the WHO–5 Well-Being Index to Identify College Students at Risk for Mental Health Problems. *Journal of College Student Development*, *58*(1), 113–117.

Guo, T., Zhao, W., Alrashoud, M., Tolba, A., Firmin, S., & Xia, F. (2022). Multimodal Educational Data Fusion for Students' Mental Health Detection. *IEEE Access*, *10*, 70370–70382. https://doi.org/10.1109/ACCESS.2022.3187502

Hellas, A., Ihantola, P., Petersen, A., Ajanovski, V. V., Gutica, M., Hynninen, T., Knutas, A., Leinonen, J., Messom, C., & Liao, S. N. (2018). Predicting academic performance: A systematic literature review. *Proceedings Companion of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education*, 175–199. https://doi.org/10.1145/3293881.3295783

Hossain, S., O'Neill, S., & Strnadová, I. (2023). What Constitutes Student Well-Being: A Scoping Review Of Students' Perspectives. *Child Indicators Research*, *16*(2), 447–483. https://doi.org/10.1007/s12187-022-09990-w

Maaliw, R. R., Quing, K. A. C., Susa, J. A. B., Marqueses, J. F. S., Lagman, A. C., Adao, R. T., Fernando - Raguro, Ma. C., & Canlas, R. B. (2022). Clustering and Classification Models For Student's Grit Detection in E-Learning. *2022 IEEE World AI IoT Congress (AIIoT)*, 039–045. https://doi.org/10.1109/AIIoT54504.2022.9817177

Pascoe, M. C., Hetrick, S. E., & Parker, A. G. (2020). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth*, *25*(1), 104–112. https://doi.org/10.1080/02673843.2019.1596823

Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2015). College Students: Mental Health Problems and Treatment Considerations. *Academic Psychiatry*, *39*(5), 503–511. https://doi.org/10.1007/s40596-014-0205-9

Rodrigues, M., Gonçalves, S., Carneiro, D., Novais, P., & Fdez-Riverola, F. (2013). Keystrokes and Clicks: Measuring Stress on E-learning Students. In J. Casillas, F. J. Martínez-López, R. Vicari, & F. De La Prieta (Eds.), *Management Intelligent Systems* (Vol. 220, pp. 119–126). Springer International Publishing. https://doi.org/10.1007/978-3-319-00569-0_15

Yao, M., Sahebi, S., & Behnagh, R. F. (2020). *Analyzing Student Procrastination in MOOCs: A Multivariate Hawkes Approach*.

Zhao, P., Li, Q., Yao, Y., & Li, Y. (2023). Precise Recognition Model for Mobile Learning Procrastination Based on Backpropagation Neural Network. *Sensors and Materials*, *35*(12), 4291. https://doi.org/10.18494/SAM4378