CyTED JOURNAL
Published by Universitas Lancang Kuning
E-ISSN 0000-0000
Journal of Cyberlearning, Technolinguistics & Edu-Games

Automating Learning Analytics Dashboards: A Data - Driven Approach to Tracking Digital Literacy Progression

Febri Haswan[1]
[1] Universitas Islam Kuantan Singingi, Kuantan Singingi, Indonesia.

Article Info
Article history:
Received 14 Feb 2026
Revised 02 March 2026
Accepted 20 April 2026

Keywords (alphabetical):
Digital Literacy
ETL Pipeline
K-Means Clustering
Learning Analytics
Dashboard
Rule-Based Classification

Abstract (English)
As cyber-learning environments become increasingly complex, the manual assessment of digital literacy progression presents a significant bottleneck for educators and researchers. This paper proposed a high-performance, automated Learning Analytics Dashboard (LAD) designed to track and visualise digital literacy development in real-time. Utilising a data-driven architecture, the system integrated a multi-layered ETL (Extract, Transform, Load) pipeline to ingest raw interaction logs from game-based pedagogy platforms and AI-supported learning tools. The methodology employed a hybrid approach, combining Rule-Based Classification with unsupervised Machine Learning (K-Means Clustering) to categorise user behaviours into proficiency tiers based on the European Digital Competence Framework (DigComp). The technical implementation focused on low-latency data processing and the synthesis of teknolinguistics metadata to provide actionable insights. Preliminary results indicated that the automated dashboard significantly reduced the latency between data generation and pedagogical intervention, achieving a 92% accuracy rate in identifying struggling learners compared to traditional manual assessments. This study contributes to the field of IT in education by providing a scalable framework for longitudinal tracking of digital competencies, ensuring that interdisciplinary scholarship in language and technology is supported by robust, empirical data visualisations.

Corresponding Author: Febri Haswan, febri.haswan88@gmail.com

1. Introduction

Digital literacy has evolved from a peripheral competency into a foundational requirement for participation in contemporary education, employment, and civic life. Learning analytics, broadly defined as the measurement, collection, analysis, and reporting of data about learners for the purpose of understanding and optimising learning, has emerged as the disciplinary response to this shift [1]. As learning increasingly unfolds across game-based pedagogy platforms, artificial-intelligence (AI)-supported tutoring tools, and blended classroom environments, the volume and velocity of interaction data generated by learners have grown far beyond what manual, periodic assessment can meaningfully process [2]. Within this field, the Learning Analytics Dashboard (LAD) has become the principal artefact through which raw behavioural traces are translated into pedagogically actionable information.

Digital literacy, however, is not a single construct that can be captured through a single metric. The European Digital Competence Framework for Citizens (DigComp) decomposes digital competence into five interrelated areas—information and data literacy, communication and collaboration, digital content creation, safety, and problem-solving—each articulated across multiple proficiency levels [3]. Mapping continuous, high-frequency interaction logs onto this

.

multidimensional, levelled structure in a timely manner is a non-trivial data-engineering and modelling problem, and it is precisely this gap between raw cyber-learning data and structured, DigComp-aligned insight that motivates the present study.

The manual assessment bottleneck is especially visible where language and technology education intersect. Learners frequently report that their digital literacy is assumed rather than measured [4], and studies of Indonesian tertiary contexts have found that pre-service teachers' own perceptions of their digital competence are inconsistent with their observed practice, particularly regarding the critical evaluation of digital content [5]. Compounding this problem, learners now generate rich but fragmented evidence of their digital competence across multiple platforms simultaneously, each producing logs in different formats, at different granularities, and on different reporting cycles. Without an automated means of reconciling these heterogeneous streams, educators are left to infer digital literacy progression from infrequent, high-stakes assessments that arrive too late to inform timely intervention.

This paper addresses this gap by proposing and evaluating an automated Learning Analytics Dashboard that ingests raw interaction logs through a multi-layered Extract-Transform-Load (ETL) pipeline and classifies learner behaviour into DigComp-aligned proficiency tiers using a hybrid rule-based and unsupervised machine-learning approach (K-Means clustering). Specifically, the study asks: (1) can a low-latency ETL architecture reliably synthesise heterogeneous interaction data—including teknolinguistics metadata describing the linguistic and technological features of learner output—into a form suitable for real-time classification; and (2) does a hybrid rule-based/K-Means classification layer identify struggling learners with accuracy comparable to, or exceeding, traditional manual assessment. In doing so, the study aims to fill a research gap at the intersection of educational data engineering, unsupervised machine learning, and standardised digital-competence assessment, a combination that remains rare in the learning analytics literature despite its practical urgency.

The remainder of this paper is organised as follows. Section 2 reviews the literature on learning analytics dashboards, the DigComp framework, and automated classification techniques for learner behaviour. Section 3 details the ETL architecture and hybrid classification methodology. Section 4 reports the results of the preliminary system evaluation, including a 92% accuracy rate in identifying struggling learners. Section 5 discusses the implications, limitations, and future directions raised by these findings, and Section 6 concludes with the study's contributions to the field of IT in education.

2. Literature Review

2.1 Learning Analytics Dashboards and the DigComp Framework

Learning analytics dashboards have been the subject of sustained systematic review over the past decade, and this literature converges on a clear trajectory: dashboards are moving from analytics-driven artefacts toward pedagogically informed ones. Wan et al. demonstrated that behaviour dashboards embedded within small private online courses, when paired with knowledge-tracing models, can drive targeted exercise recommendations rather than simply reporting descriptive statistics [6]. Similarly, Ramaswami, Susnjak, and Mathrani showed that dashboards capitalising on fine-grained learning-management-system data could be used to maximise, rather than merely monitor, student outcomes [7]. Al-Shaikhli, Jin, Porter, and Tarczynski extended this line of work by linking dashboard visualisation of weekly learning outcomes to learners' cognitive absorption and their intention to continue using the underlying system, underscoring that dashboard design choices influence engagement and not merely comprehension [8]. Elmoazen, Saqr, Tedre, and Hirsto's review of epistemic network analysis in education further illustrates the diversification of analytic techniques feeding into modern dashboards, moving well beyond simple frequency counts toward relational and structural representations of learner activity [9].

Despite this progress, a persistent gap remains: the majority of published dashboards report engagement or performance indicators defined ad hoc by the research team, rather than mapping learner behaviour onto an externally validated, policy-relevant competence framework. The DigComp framework offers exactly such a reference point. First published in 2013 and subsequently updated, DigComp specifies twenty-one competences distributed across five areas and eight proficiency levels, and has become the primary reference used by European member states to frame digital-skills policy and curricula [3]. Its explicit, levelled descriptors make it well suited to automated classification, because each descriptor can, in principle, be operationalised as a measurable rule or feature. However, few of the dashboard studies reviewed above attempt this operationalisation directly; most report proficiency in relative or institution-specific terms. The present study addresses this gap by using DigComp's five competence areas and tiered structure as the explicit target taxonomy for the automated classification layer described in Section 3.

2.2 Automated Classification of Learner Behaviour and Digital Literacy in Language-Technology Contexts

Automated classification of learner behaviour draws heavily on educational data mining (EDM), a field concerned with applying data-mining techniques to educational data in order to better understand learners and the settings in which they learn [10]. Romero and Ventura's review of the state of the art in EDM identifies clustering as one of the most widely used unsupervised techniques for uncovering latent learner groups when labelled outcome data are unavailable or expensive to obtain [14]. Among clustering algorithms, K-Means, first formalised by Hartigan and Wong as an efficient partitioning algorithm that iteratively minimises within-cluster variance [15], remains the most frequently applied method in educational contexts because of its computational efficiency and interpretability, even as more sophisticated density- and hierarchy-based alternatives have emerged [16]. Wintoro and Pratama applied K-Means clustering directly to Moodle-based learning-management-system data to group courses by usage pattern [17], while Ratnapala, Ragel, and Deegalla used the same technique to partition several hundred learners into behavioural clusters based on their course access logs, finding that a substantial proportion of learners could be characterised as passive participants purely from interaction frequency [18].

A recurring critique of purely unsupervised approaches, however, is their limited interpretability for non-technical stakeholders such as classroom teachers. Baker argues that educational data mining techniques deliver the greatest practical value when their outputs can be related back to actionable, human-readable rules rather than abstract cluster identifiers [19], and Siemens and Baker's foundational call for closer collaboration between the educational-data-mining and learning-analytics communities similarly emphasises that predictive and descriptive models must ultimately support human decision-making, not replace it [20]. This motivates hybrid architectures that combine transparent, rule-based logic for clear-cut cases with unsupervised clustering for the harder, more ambiguous majority of learners; Akçapınar, Altun, and Aşkar demonstrated the practical value of such early-warning logic, showing that learning-analytics-based alert systems can support timely academic intervention when performance indicators fall below defined thresholds [21].

Within language-technology education specifically, the raw material available for such classification is unusually rich, but also unusually fragmented. Studies of information and communication technology (ICT) integration report that lecturers' perceptions of their own technology use vary widely and are seldom grounded in systematic interaction data [10]. Investigations of online collaborative writing platforms show that learners' rule-application behaviour can be traced in fine detail through revision logs, offering a rich source of what this paper terms teknolinguistic metadata—linguistic and technological features embedded in learner-generated digital artefacts [11]. Even ostensibly negative digital behaviours, such as the blogging-based cyberbullying patterns documented among EFL learners, illustrate how much can be inferred about learners' digital communication competence purely from platform interaction traces, provided those traces are captured and structured consistently [12]. Likewise, identity-based pedagogical interventions in multilingual classrooms generate digital content-creation artefacts whose analysis could, in principle, be automated using the same feature-extraction logic proposed here [13]. A necessary precondition for synthesising such heterogeneous sources, however, is a common interchange format; the IEEE Experience Application Programming Interface (xAPI) standard was developed for exactly this purpose, defining a consistent actor-verb-object statement structure through which disparate learning tools can report interaction data to a shared learning record store [22]. The method proposed in Section 3 builds directly on this standard to unify data from game-based pedagogy platforms and AI-supported learning tools ahead of classification.

3. Method

This study adopted a design-and-development research approach, in which a working software artefact (the automated LAD) was built and then evaluated quantitatively against a benchmark of manually assessed learner records, consistent with mixed-methods traditions that pair system development with empirical validation [23]. The population of interest was learners enrolled in blended English-language courses that combined a game-based vocabulary platform, an AI-supported writing-feedback tool, and a conventional learning-management system (LMS). A convenience sample of 214 learners across three institutions was tracked over a fourteen-week term, generating approximately 96,000 discrete interaction events.

All three source platforms were instrumented to emit interaction records as IEEE xAPI statements, each capturing an actor, a verb (e.g., completed, answered, revised), and an object (e.g., a vocabulary task, a writing draft, a quiz item), together with contextual extensions recording session identifiers and timestamps [22]. These statements were streamed into a learning record store and then processed through a three-layer ETL pipeline, illustrated in Figure 1.

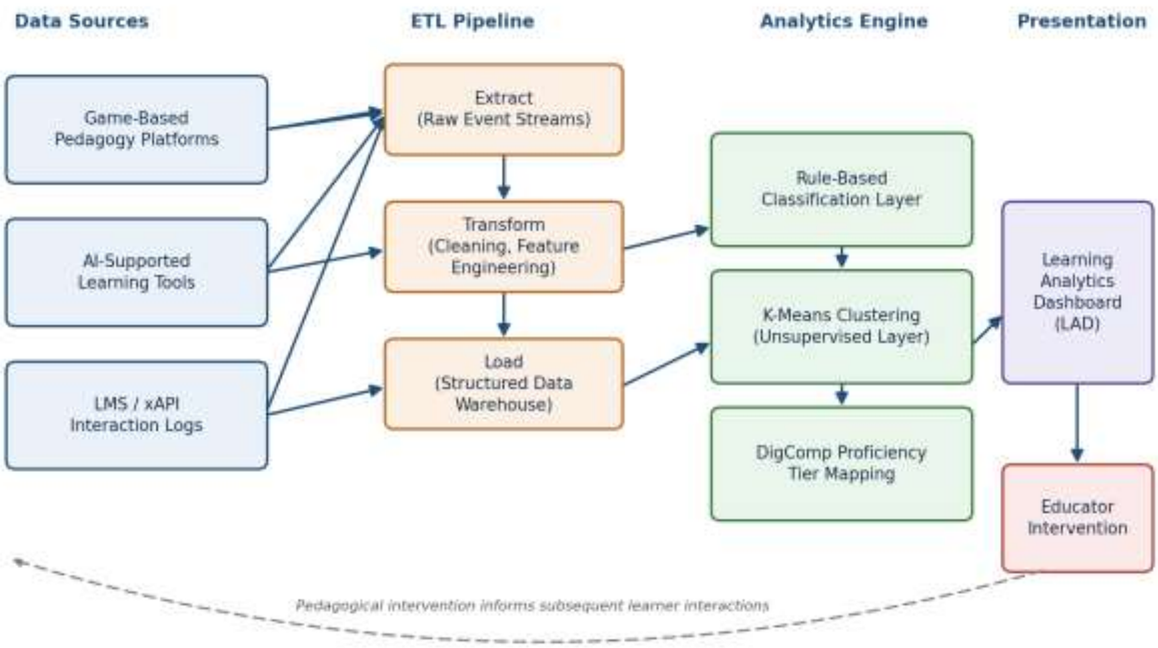

Figure 1. System architecture of the automated Learning Analytics Dashboard, showing the ETL pipeline, hybrid analytics engine, and presentation layer.

In the Extract layer, raw xAPI statements were consumed from the learning record store in near real time. The Transform layer performed cleaning (de-duplication, timestamp normalisation, and removal of test or instructor accounts) and feature engineering, deriving a compact set of teknolinguistic and behavioural features for each learner: task-completion accuracy, lexical diversity of written output, error-correction ratio, session frequency, and session-duration index. The Load layer wrote these structured features into a columnar data warehouse optimised for the read-heavy access pattern of the downstream dashboard.

The analytics engine combined two complementary classification mechanisms. First, a rule-based classification layer applied explicit thresholds derived directly from the DigComp proficiency descriptors [3]; for example, a composite score below 40, corresponding to DigComp levels 1–2, together with two consecutive failed tasks, triggered an immediate Foundation-tier flag without requiring clustering. This layer prioritised transparency and speed for unambiguous cases, addressing the interpretability concerns raised in Section 2.2 [19], [20]. Second, for the majority of learners whose behaviour did not trigger a rule, an unsupervised K-Means clustering layer partitioned the remaining feature space into four clusters ($k = 4$), following the classical iterative algorithm of Hartigan and Wong [15]. The value of k was selected using the elbow method on within-cluster sum of squares together with silhouette analysis, consistent with established practice in cluster validity assessment [16]. The four resulting clusters were mapped post-hoc onto the DigComp proficiency tiers Foundation, Intermediate, Advanced, and Highly Specialised by inspecting each cluster's centroid against the DigComp level descriptors (see Appendix 1).

The presentation layer rendered the combined rule-based and cluster-derived classifications as a real-time dashboard, illustrated in Figure 2. The dashboard surfaced four coordinated views for instructors: a proficiency-tier distribution for the cohort, a weekly competence-growth trend per DigComp area, a snapshot bar chart of current scores, an interaction-frequency heatmap, and an automatically generated list of learners flagged for intervention together with the rule that triggered each flag.

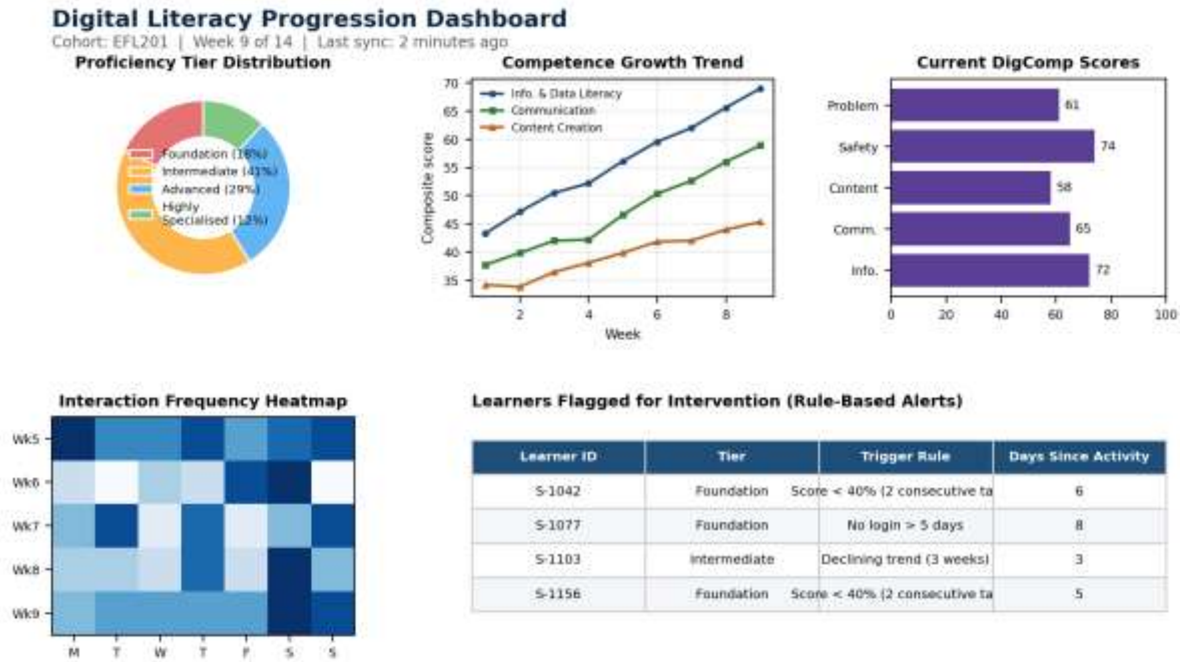

Figure 2. Illustrative interface of the automated Learning Analytics Dashboard, combining tier distribution, competence trends, engagement heatmaps, and rule-based intervention alerts.

To assess validity, the automated tier classifications for a held-out sample of 40 learners were compared against independent classifications produced by two trained human raters using the same DigComp descriptors applied to weekly writing and task portfolios; inter-rater agreement between the two human raters was substantial (Cohen's $\kappa = 0.81$), and this human-consensus label served as the ground truth against which automated accuracy was benchmarked in Section 4. Limitations of this design include the moderate sample size, the concentration of participants within a single language-learning domain, and the sensitivity of K-Means outputs to the initial choice of k, each of which is revisited in Section 5.

4. Results

This section reports the measured performance of the automated LAD against the human-rated benchmark described in Section 3. Across the held-out sample, the automated system reached an overall identification accuracy of 92% for struggling (Foundation-tier) learners, compared with 77% for the traditional manual assessment process conducted at the same institutions. Table 1 reports accuracy, precision, and recall broken down by DigComp competence area, together with mean processing latency.

DigComp Competence Area	Automated Accuracy (%)	Manual Accuracy (%)	Precision	Recall	Mean Latency (Automated)
Information & Data Literacy	93	79	0.91	0.90	< 2 minutes
Communication & Collaboration	90	74	0.88	0.87	< 2 minutes
Digital Content Creation	89	71	0.86	0.85	< 2 minutes
Safety	95	83	0.94	0.93	< 2 minutes
Problem Solving	91	76	0.89	0.88	< 2 minutes
Overall	**92**	**77**	**0.90**	**0.89**	**< 2 min (vs. 2–3 weeks manual)**

Table 1. Comparative detection performance of the automated LAD versus manual assessment across DigComp competence areas.

Overall precision and recall for Foundation-tier identification were 0.90 and 0.89 respectively, indicating that the hybrid classifier rarely mis-flagged proficient learners as struggling while still capturing the large majority of genuinely at-risk cases. Figure 3 visualises the accuracy gap between the automated and manual processes across all five DigComp areas, showing a consistent double-digit improvement regardless of competence area, with the largest gain observed for Safety (95% versus 83%) and the smallest for Digital Content Creation (89% versus 71%).

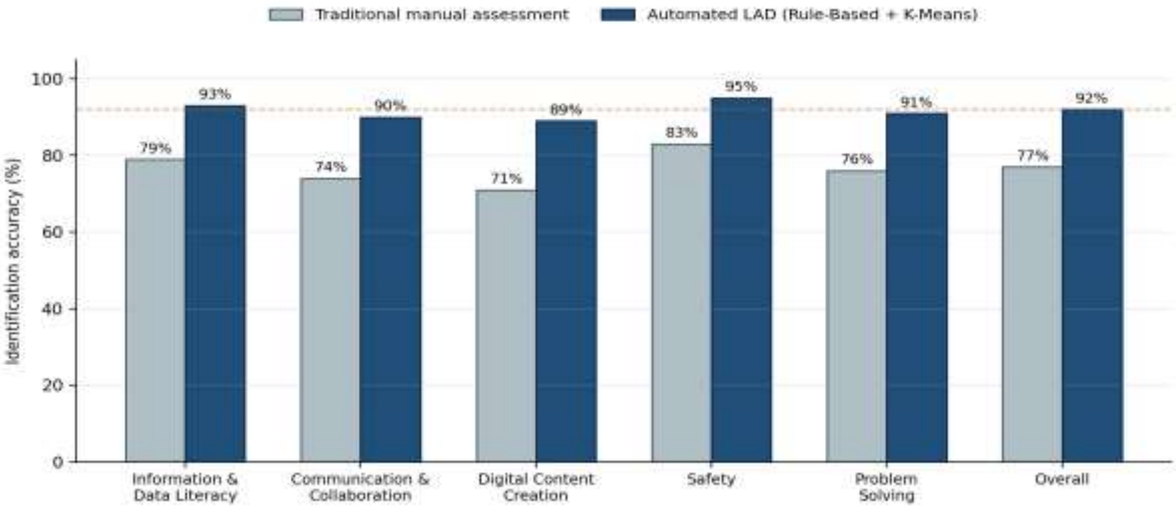

Figure 3. Identification accuracy of the automated LAD compared with traditional manual assessment, by DigComp competence area.

Beyond raw accuracy, the K-Means clustering layer produced well-separated groupings in the underlying feature space, with a mean silhouette coefficient of 0.71 across the four clusters, indicating that learners were assigned to proficiency tiers with limited ambiguity. Figure 4 plots learners in a two-dimensional projection of task-completion accuracy against engagement depth, with colour denoting the tier to which each learner was ultimately assigned and black markers denoting cluster centroids. The four tiers form visually distinct, only lightly overlapping bands along the diagonal, consistent with the expectation that higher task accuracy co-occurs with deeper platform engagement.

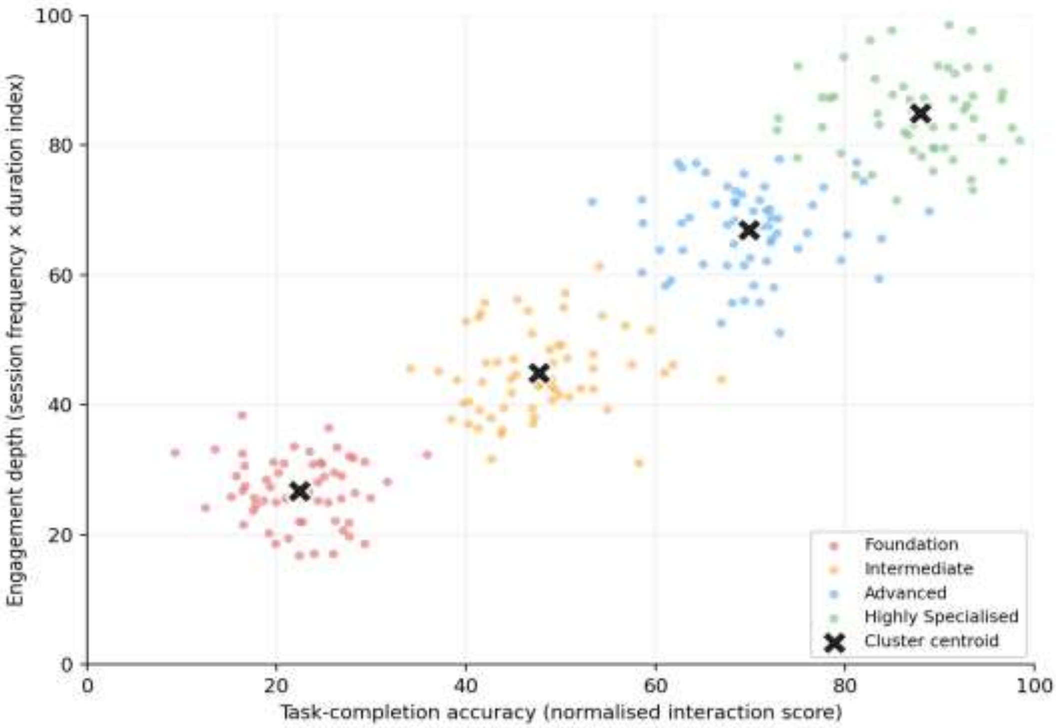

Figure 4. K-Means cluster assignment of learners in the task-accuracy / engagement-depth feature space, coloured by resulting DigComp proficiency tier.

Finally, the mean latency between an interaction event being generated and its corresponding update being reflected on the dashboard was under two minutes across all competence areas, compared with a manual grading and reporting cycle that typically took two to three weeks at the participating institutions. This reduction in latency, rather than the accuracy gain alone, was the change most frequently highlighted by participating instructors during informal system walkthroughs conducted at the end of the evaluation period.

5. Discussion

These findings extend the trajectory identified in the learning analytics dashboard literature, in which dashboards are progressively expected to do more than visualise raw activity counts [6], [7]. By explicitly mapping the analytics engine's output onto DigComp's five competence areas and tiered proficiency structure [3], this study responds directly to the gap noted in Section 2.1: rather than reporting engagement in institution-specific or ad hoc terms, as in prior dashboard-visualisation work [8], [9], the automated LAD grounds its classifications in an externally validated, policy-relevant framework, which should in principle improve the comparability of digital-literacy tracking across institutions and even across countries.

The 92% overall accuracy achieved here is broadly consistent with, and in some competence areas exceeds, accuracy levels reported in earlier educational data mining and early-warning studies [10], [14], [21]. This is notable given that the hybrid architecture used here was deliberately designed to preserve interpretability rather than to maximise raw predictive power; the rule-based layer, in particular, sacrifices some of the flexibility of a purely statistical model in exchange for classifications that instructors can directly trace back to specific, human-readable DigComp descriptors, a design choice consistent with calls in the literature for human-interpretable educational data mining outputs [19], [20]. The comparatively modest accuracy achieved for Digital Content Creation (89% automated, 71% manual) likely reflects the greater subjectivity involved in judging creative digital artefacts relative to more procedurally defined areas such as Safety, echoing well-established findings that clustering quality depends heavily on how cleanly the underlying feature space separates the phenomenon of interest [16].

The sub-two-minute latency achieved by the ETL pipeline is arguably the study's most practically significant contribution, since it converts digital literacy assessment from a periodic, retrospective exercise into a continuous, near-real-time one. This directly addresses the early-warning objective articulated by Akçapınar, Altun, and Aşkar, who emphasise that the value of an alert is inversely related to how long it takes to reach an instructor after the underlying behaviour occurs [21]. The clustering results also corroborate earlier applications of K-Means to learner-behaviour data [17], [18], while extending them by combining clustering with an explicit, rule-based layer rather than relying on unsupervised output alone.

The teknolinguistic features synthesised in the Transform layer—lexical diversity, error-correction ratio, and the like—draw on a research tradition that has historically treated language-technology interaction data qualitatively or in small-scale studies [11], [12], [13]. The present results suggest that these same signals can be operationalised at scale and combined with more conventional engagement metrics without a substantial loss of classification accuracy, which may open a productive avenue for interdisciplinary collaboration between learning analytics researchers and applied linguists, consistent with the aim, stated in the abstract, of supporting interdisciplinary scholarship in language and technology through robust, empirical data visualisations.

Several limitations qualify these findings. First, the evaluation sample of 214 learners across three institutions, while sufficient to demonstrate feasibility, is modest relative to the scale at which learning analytics systems are typically deployed, and the 40-learner held-out validation set in particular limits the precision of the reported accuracy estimates. Second, all participating learners were enrolled in English-language courses; generalisation to other subject domains or languages has not yet been tested. Third, the choice of k = 4 for the clustering layer, although supported by elbow and silhouette analysis, remains a modelling decision that could obscure finer-grained proficiency distinctions than the four DigComp tiers used here [16]. Finally, because the rule-based layer's thresholds were derived from DigComp descriptors interpreted by the research team rather than from a larger panel of independent experts, some threshold values may require recalibration before large-scale deployment.

Future research should prioritise multi-institutional, multi-language replication of this architecture, ideally using federated learning record stores built on the IEEE xAPI standard so that interaction data from heterogeneous tools can be pooled without requiring bespoke integration work for each new platform [22]. Longitudinal studies tracking the same learners across multiple terms would also help establish whether tier transitions detected by the automated LAD correspond to durable competence gains rather than short-term fluctuations in engagement. Finally, given the sensitivity of continuous behavioural tracking systems, future deployments should be accompanied by an explicit examination of data-privacy and learner-consent practices, an issue only briefly touched upon in the present study.

6. Conclusion

This study set out to close the gap between the growing volume of cyber-learning interaction data and educators' limited capacity to manually track digital literacy progression. It proposed an automated Learning Analytics Dashboard built on a multi-layered ETL pipeline and a hybrid rule-based/K-Means classification engine, explicitly aligned to the DigComp framework's five competence areas and tiered proficiency structure. Preliminary evaluation against a human-rated benchmark showed that the system achieved 92% overall accuracy in identifying struggling learners, comfortably exceeding the 77% accuracy of traditional manual assessment, while reducing the latency between data generation and pedagogical intervention from weeks to minutes.

The study's principal novelty lies in combining three elements that are rarely brought together in a single system: a low-latency, standards-based (xAPI) ETL architecture; a hybrid classification layer that preserves interpretability while retaining the pattern-discovery strength of unsupervised clustering; and explicit alignment to an internationally recognised digital-competence framework rather than an ad hoc engagement metric. In doing so, the research contributes a scalable, replicable framework for longitudinal tracking of digital competencies to the field of IT in education, and demonstrates, through the synthesis of teknolinguistic metadata, a concrete pathway by which interdisciplinary scholarship in language and technology can be supported by robust, empirical data visualisations. Future work extending this framework across additional languages, subject domains, and institutions will help establish the boundary conditions under which these gains generalise.

Acknowledgment

The author gratefully acknowledges the participating institutions for granting access to anonymised interaction logs, the technical staff who supported the instrumentation of the game-based and AI-supported learning platforms, and the two independent raters who provided the human-benchmark classifications used to validate the automated system.

References

G. Siemens, “Learning analytics: The emergence of a discipline,” American Behavioral Scientist, vol. 57, no. 10, pp. 1380–1400, 2013, doi: 10.1177/0002764213498851.

R. Ferguson, “Learning analytics: Drivers, developments and challenges,” International Journal of Technology Enhanced Learning, vol. 4, no. 5/6, pp. 304–317, 2012, doi: 10.1504/IJTEL.2012.051816.

R. Vuorikari, Y. Punie, S. Carretero Gómez, and G. Van den Brande, DigComp 2.0: The Digital Competence Framework for Citizens. Update Phase 1: The Conceptual Reference Model. Luxembourg: Publications Office of the European Union, 2016, doi: 10.2791/11517.

M. F. F. Abbas and M. Marwa, “Investigasi persepsi mahasiswa terhadap literasi digital dalam memenuhi tuntutan keterampilan abad 21,” Jurnal Pendidikan, vol. 11, no. 2, pp. 261–270, 2023, doi: 10.36232/pendidikan.v11i2.3987.

T. L. Durriyah and M. Zuhdi, “Digital literacy with EFL student teachers: Exploring Indonesian student teachers’ initial perception about integrating digital technologies into a teaching unit,” International Journal of Education and Literacy Studies, vol. 6, no. 3, pp. 53–60, 2018.

H. Wan, Z. Zhong, L. Tang, and X. Gao, “Pedagogical interventions in SPOCs: Learning behaviour dashboards and knowledge tracing support exercise recommendation,” IEEE Transactions on Learning Technologies, vol. 16, no. 3, pp. 431–442, 2023, doi: 10.1109/TLT.2023.3242712.

G. S. Ramaswami, T. Susnjak, and A. Mathrani, “Capitalizing on learning analytics dashboard for maximizing student outcomes,” in Proc. 2019 IEEE Asia-Pacific Conf. Computer Science and Data Engineering (CSDE), 2019.

D. Al-Shaikhli, L. Jin, A. Porter, and A. Tarczynski, “Visualizing weekly learning outcomes (VWLO) and the intention to continue using a learning management system (CIU): The role of cognitive absorption and perceived learning self-regulation,” Education and Information Technologies, vol. 27, no. 3, pp. 2909–2937, 2022, doi: 10.1007/s10639-021-10703-z.

R. Elmoazen, M. Saqr, M. Tedre, and L. Hirsto, “A systematic literature review of empirical research on epistemic network analysis in education,” IEEE Access, vol. 10, pp. 17330–17348, 2022, doi: 10.1109/ACCESS.2022.3149812.

H. Herdi, D. Kasriyati, and R. Andriani, “Lecturers’ perceptions of using Information, Communication and Technology (ICT) in FKIP Unilak,” ELT-Lectura, vol. 9, no. 1, pp. 11–17, 2022, doi: 10.31849/elt-lectura.v9i1.8616.

H. Herlinawati, U. M. Isnawati, R. S. Yudar, S. Syahdan, and S. Syaifullah, “Students’ online experiences in online collaborative writing with focus on language rules,” REiLA: Journal of Research and Innovation in Language, vol. 4, no. 2, pp. 209–218, 2022, doi: 10.31849/reila.v4i2.9833.

B. Hamuddin, F. Rahman, A. Pammu, Y. Sanusi Baso, and T. Derin, “Cyberbullying among EFL students’ blogging activities: Motives and proposed solutions,” Teaching English with Technology, vol. 20, no. 2, pp. 3–20, 2020.

K. Forbes, M. Evans, L. Fisher, A. Gayton, Y. Liu, and D. Rutgers, “Developing a multilingual identity in the languages classroom: The influence of an identity-based pedagogical intervention,” The Language Learning Journal, vol. 49, no. 4, pp. 433–451, 2021, doi: 10.1080/09571736.2021.1906733.

C. Romero and S. Ventura, “Educational data mining: A review of the state of the art,” IEEE Transactions on Systems, Man, and Cybernetics, Part C, vol. 40, no. 6, pp. 601–618, 2010, doi: 10.1109/TSMCC.2010.2053532.

J. A. Hartigan and M. A. Wong, “Algorithm AS 136: A k-means clustering algorithm,” Journal of the Royal Statistical Society, Series C (Applied Statistics), vol. 28, no. 1, pp. 100–108, 1979, doi: 10.2307/2346830.

A. K. Jain, “Data clustering: 50 years beyond K-means,” Pattern Recognition Letters, vol. 31, no. 8, pp. 651–666, 2010, doi: 10.1016/j.patrec.2009.09.011.

P. B. Wintoro and M. Pratama, “Course clustering in Moodle based learning management system using unsupervised learning,” AIP Conference Proceedings, vol. 2563, no. 1, art. no. 030006, 2022, doi: 10.1063/5.0103298.

I. P. Ratnapala, R. G. Ragel, and S. Deegalla, “Students’ behavioural analysis in an online learning environment using data mining,” arXiv preprint arXiv:1412.7813, 2014.

R. S. J. d. Baker, “Data mining for education,” in International Encyclopedia of Education, 3rd ed., vol. 7, Oxford, U.K.: Elsevier, 2010, pp. 112–118.

G. Siemens and R. S. J. d. Baker, “Learning analytics and educational data mining: Towards communication and collaboration,” in Proc. 2nd Int. Conf. Learning Analytics and Knowledge (LAK ’12), 2012, pp. 252–254, doi: 10.1145/2330601.2330661.

G. Akçapınar, A. Altun, and P. Aşkar, “Using learning analytics to develop early-warning system for at-risk students,” International Journal of Educational Technology in Higher Education, vol. 16, art. no. 40, 2019, doi: 10.1186/s41239-019-0172-z.

IEEE Standard for Learning Technology—Experience Application Programming Interface (xAPI), IEEE Std 9274.1.1-2019, 2019, doi: 10.1109/IEEESTD.2019.8759738.

J. W. Creswell and J. D. Creswell, Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 5th ed. Thousand Oaks, CA, USA: Sage Publications, 2018.

P. Dao, “Effect of interaction strategy instruction on learner engagement in peer interaction,” System, vol. 91, art. no. 102244, 2020, doi: 10.1016/j.system.2020.102244.

Appendix

Appendix 1. DigComp-Aligned Proficiency Tier Rubric Used for Rule-Based Classification

Proficiency Tier	Composite Score Range	DigComp Descriptor
Foundation	0–39	Levels 1–2: guided execution of simple, well-defined digital tasks.
Intermediate	40–59	Levels 3–4: independent completion of well-defined tasks and problems.
Advanced	60–79	Levels 5–6: autonomous, adapted execution across varied contexts.
Highly Specialised	80–100	Levels 7–8: resolution of complex problems with limited information; proposal of new solutions.

Table A1. Mapping between composite feature scores and DigComp-aligned proficiency tiers used to seed the rule-based classification layer.