Designing a multidimensional profiling framework for English learners

SUCI MADIARTI ISMAN[1*], IZZUL FATAWI[2], AND DWI REZKI HARDIANTO PUTRA RUSTAN[3]

Abstract

This research presents the design of a multidimensional profiling framework to support adaptive and personalized English language learning at Universitas Terbuka (UT). A qualitative research method was used in designing the profiling framework. Data were collected through focus group discussions (FGDs) and institutional observations to examine the current realities within the university. Thematic analysis was then employed to interpret the data. Two key results were a learner profiling framework and an institutional readiness mapping. The framework included four phases: (1) Input: collecting data on learners' digital readiness, learning styles, English proficiency, and self-regulated learning; (2) Processing: analyzing the data to create learner profiles; (3) Adaptation: using profiles to suggest personalized support; and (4) Interpretation: presenting results through dashboards and feedback tools. The institutional mapping found that each phase required a clear governance structure to ensure ethical use, transparency, and alignment with Universitas Terbuka's policies and stakeholder capacities.

Keywords

English language, framework design, learner profiling, learning analytics, Universitas Terbuka (UT)

Article History

Received 01 April 2025
Accepted 27 August 2025

How to Cite

Isman, S. M., Fatawi, I., & Rustan, D. R. H. (2025). Designing a multidimensional profiling framework for English learners. *Indonesian Research Journal in Education |IRJE|*, *9*(2), 1126 - 1141. https://doi.org/10.22437/irje.v9i02.47643

[1] Universitas Terbuka, Banten, Indonesia, Corresponding author: sisman@ecampus.ut.ac.id.
[2] Universitas Terbuka, Banten, Indonesia
[3] Universitas Terbuka, Banten, Indonesia

Introduction

In the evolving landscape of open and distance learning (ODL), the personalization of instruction has emerged as a critical success factor, particularly in large-scale institutions such as Universitas Terbuka (UT), Indonesia, which currently serves over 650,000 students across a diverse archipelago. With the growth of data availability and technological affordances, learning analytics has increasingly been used to support individualized support systems, including student profiling (Viberg et al., 2018). In language learning contexts, accurate learner profiling is crucial for addressing diverse proficiency levels and supporting motivation, self-regulation, and engagement (Guo et al., 2025).

Despite the potential of profiling systems utilization to improve educational outcomes, few frameworks have been adapted to the specific constraints and affordances, as well as the contextual aspects of ODL environments. Nonetheless, the literature contains several studies conducted on student profiling in the ODL context. Carow et al. (2023) had profiled the students' lives and demographics in South Africa to examine their learning readiness within an Open and Distance e-Learning (ODeL) environment. The findings revealed that students living in urban areas and accessing internet hotspots experienced greater ease in navigating the ODeL environments. In contrast, those in rural areas faced the most significant challenges.

Furthermore, a study by Kuruppuarachchi and Karunanayake (2017) explored the socio-economic conditions of BSc undergraduates at the Open University of Sri Lanka (OUSL) and their understanding of the ODL concept. Finally, Latif et al. (2020) did a study that focused more extensively on demographic profiles, personality, attitudes, and motivation to address the high dropout rates in the first semesters at the Open University Malaysia (OUM). Their findings indicated a strong correlation between students' motivation, educational goals, career aspirations, and academic achievement. These exploratory studies focused on learner profiling without establishing a contextual design framework within their institutions.

Ozturk et al. (2019) proposed a mapping model for a learner profiling system. However, several critical learner dimensions, such as digital readiness, learning styles, and self-regulation, were not fully integrated. Another study on profiling English learners found that higher academic achievement is strongly associated with higher English proficiency, greater readiness for independent learning, and access to better digital resources. While educational background and teaching experience played a role, they were less conclusive (Isman et al., 2024).

At Universitas Terbuka (UT), institutional data such as demographic, academic history, and online learning behavior are available. They can be comprehensively integrated into a profiling system that provides actionable feedback for students or instructors. Moreover, the critical learner dimensions necessary in student profiling, such as digital readiness, learning styles, and self-regulation, can also be assessed systematically. Considering the findings of the cited studies, the potential availability of data on student learning, and the need to enhance the online learning experience of students through a more personalized learning path, a study to design a multidimensional profiling framework for English learners at Universitas Terbuka (UT) was carried out.

Literature Review

Student profiling in distance and online education

Student profiling in higher education involves systematically collecting and analyzing learner data to support personalized learning, retention, and academic success. In online and distance learning (ODL) contexts, where face-to-face interactions are minimal and learner diversity is high, profiling is essential for tailoring instruction and interventions (Ifenthaler & Yau, 2020; Viberg et al., 2018).

Open and distance learning institutions like Universitas Terbuka (UT) typically serve non-traditional students, including working professionals, adult learners, and individuals from geographically remote or underserved areas. This heterogeneity demands a multidimensional profiling approach, integrating demographic data with cognitive, behavioral, and affective dimensions (Ifenthaler & Yau, 2020; Viberg et al., 2018). Another study has shown that key variables influencing academic performance and persistence in online education include self-regulation, digital literacy, motivation, learning styles, and readiness for independent learning (Tsai et al., 2020). Recent findings by Ma and She (2024) underscored the motivational foundations of self-regulated learning (SRL) in online environments. This research demonstrated that self-efficacy, engagement, or satisfaction did not mediate the positive relationship between learning goal orientation and academic achievement. However, academic self-efficacy and learning engagement did mediate the effect of goal orientation on learning satisfaction, highlighting the complicated relationship between motivation and learner outcomes in digital contexts.

Recent advancements in learning analytics and educational data mining enable institutions to detect patterns in student engagement and performance data, helping identify at-risk learners, inform feedback systems, and facilitate adaptive learning pathways (Al-Zahrani & Alasmari, 2023; Wise & Vytasek, 2017). However, profiling systems must be designed with ethical considerations in mind. Researchers caution against over-reliance on predictive models without adequate regard for student agency, privacy, and institutional responsibility (Wise et al., 2021). Ethical implementation includes transparency, consent, and the responsible interpretation of data, particularly in high-stakes contexts like language learning.

At Universitas Terbuka (UT), where most English learners are in-service teachers balancing professional duties with study, profiling systems must be scalable, ethically grounded, and aligned with pedagogical goals. The literature highlights the importance of ensuring that such systems support academic personalization and fit within the technological and institutional infrastructure of ODL environments (Ifenthaler & Yau, 2020; Viberg et al., 2018).

Frameworks in learning analytics and student modelling

Learning analytics (LA) and educational data mining (EDM) have matured as fields of research and practice, leading to conceptual frameworks that guide institutions in integrating student data into teaching, learning, and decision-making. These frameworks serve as blueprints for designing systems that analyze learner behavior, support personalization, and ensure ethical implementation at scale (Viberg et al., 2018; Wise & Vytasek, 2017).

A prominent model updated for contemporary use was presented by Ifenthaler and Yau (2020), who proposed an integrated system architecture featuring three interdependent engines: the profiling engine, the analytics engine, and the personalization engine. This framework enabled fine-grained learner modelling and dynamic feedback, making it well-suited for adaptive learning systems. Its emphasis on functionality, scalability, and learning-centered design was aligned with the practical demands of large-scale open and distance education institutions.

Similarly, Al-Zahrani and Alasmari (2023) offered a multidimensional framework that addresses learning analytics systems' input, process, and output layers. Their model emphasized the systematic integration of learner attributes (e.g., cognitive, behavioral, contextual) and underscored the importance of feedback loops through dashboards and visualizations. Their design facilitated timely decision-making for instructors, administrators, and learners, making it suitable for decentralized environments like UT.

Other researchers, such as Viberg et al. (2018), have proposed frameworks incorporating institutional strategy and operational components. Their approach included policy, infrastructure, and capacity-building dimensions, offering a comprehensive lens through which institutions can assess their readiness for LA adoption. Of relevance is their emphasis on aligning LA implementation with pedagogical intent and ethical considerations, especially in culturally diverse and distributed learner populations.

Although earlier models by Siemens and Long (2011) or Greller and Drachsler (2012) have laid the foundation for ethical and functional aspects of learning analytics, more recent frameworks have built upon these ideas to account for emerging technologies, data governance practices, and learner agency (Wise et al., 2021). In designing the Universitas Terbuka (UT) framework, this research adopted the strengths of the updated model, emphasizing system integration, ethical alignment, and pedagogical relevance, while tailoring them to the specific challenges of large-scale distance English education programs in Indonesia.

Profiling dimensions for language learners

In open and distance education environments, profiling language learners requires identifying multidimensional variables influencing learning engagement, language acquisition, and academic performance in the context of Universitas Terbuka's students, especially those enrolled in the English Education Department. The most relevant core dimensions to focus on in this research are learning styles, digital readiness, English proficiency, and readiness for independent learning. These dimensions are particularly relevant in self-paced and asynchronous settings, where learners must navigate content independently with limited synchronous support.

Learning styles refer to students' preferred modes of absorbing and processing information. While early models like VARK have been widely referenced, recent research emphasizes a more dynamic interpretation of learning preferences, recognizing their interaction with context and task type. These learning preferences were shaped by integrating perceptual and cognitive styles, which influence or support material design (Camarillas, 2019) and learner engagement (du Plooy et al., 2024). Profiling learners based on style preferences helps instructors design multimodal resources and better scaffold learning pathways.

Digital readiness encompasses access to devices and connectivity, and learners' digital literacy and confidence in navigating online learning systems. Lin and Tsou (2025) highlighted the increasing importance of digital self-efficacy and navigation fluency in determining learner success in online courses. Inadequate digital readiness can lead to barriers in participation, self-regulation, and access to academic support, particularly relevant in a geographically dispersed student population like UT's.

English proficiency remains a central profiling variable, especially in English as a Foreign Language (EFL) programs. Proficiency across core skills such as listening, speaking, reading, and writing affects how students interact with learning materials, complete assessments, and engage in academic discourse. Guo et al. (2025) found that motivation and language proficiency jointly predict EFL learners' online participation and persistence. Diagnostic assessment of proficiency enables appropriate scaffolding and placement, enhancing learning alignment and reducing cognitive overload.

Readiness for independent learning relates to learners' capacity for self-direction, time management, goal-setting, and reflective thinking. In the ODL context, this dimension was closely tied to persistence and completion rates. Pintrich (1999) stressed the role of metacognition and motivational regulation in sustaining self-regulated learning. Profiling learners' readiness allowed institutions to identify those who may benefit from early interventions such as study skills modules, peer support, or guided tutoring.

Considering the cited literature's findings, these four dimensions-learning styles, digital readiness, English proficiency, and readiness for independent learning-form a robust and practical foundation for learner profiling in Universitas Terbuka's English language programs. Identifying gaps across these variables allows the institution to implement responsive support mechanisms, optimize learning design, and ensure a more personalized and equitable learning experience.

Methodology

Research design, site, and data source

This research adopted a qualitative method, employing a constructivist approach as Creswell (2014) outlined. Within this research design, the researchers aimed to interpret empirical phenomena such as data systems, institutional practices, learning environments, and the interpretations of institutional policymakers and other relevant stakeholders in constructing a profiling framework for English learners within an ODL context. In other words, the foundation of the profiling framework was constructed from the empirical realities observed at Universitas Terbuka (UT).

The research was conducted within the English Education Department (EED) at Universitas Terbuka (UT), with an enrollment of over 1,600 undergraduates across Indonesia. These students were predominantly in-service English teachers. This research drew on three primary data sources, such as (1) institutional documents; (2) existing online learning systems currently used at Universitas Terbuka (UT) to identify the conceptual and architectural models applied in profiling and learning analytics systems; and (3) key expert participants, comprising experts in educational technologies, IT developers, course coordinators, and tutors involved in Universitas Terbuka's digital education infrastructure. Meanwhile, secondary data were obtained from journal articles published within the last ten years from the Scopus, ERIC, Web of Science, and Google Scholar databases and three selected frameworks as comparative data.

Data collection

The frameworks that were examined and included in this research were selected based on their relevance to online or distance learning environments, inclusion of student profiling, analytics, and personalized feedback mechanisms, integration of learner characteristics, system data, and adaptive technologies, recognition in peer-reviewed literature or widespread academic use, and compatibility with Universitas Terbuka's institutional goals and digital infrastructure.

Based on the nature of the data, two techniques were employed to collect information from the primary data sources, such as observation and Focus Group Discussion (FGD). Observations were used to identify the conceptual and architectural models embedded in the profiling and learning analytics systems within Universitas Terbuka's Moodle-based Learning Management System (LMS). Meanwhile, FGDs were conducted to deliberate on the design of the multidimensional profiling framework. These discussions also served as a platform to ensure the robustness and contextual relevance of the insights gathered. Finally, the outcomes of these data collection techniques were triangulated with the secondary data previously reviewed and analyzed.

Data analysis

The collected data were analyzed through Braun and Clarke's (2013) thematic analysis, complemented by framework analysis to ensure depth and structure in theme development. The analysis followed a five-step procedure applied to institutional documents and observation data: (1) the researchers familiarized themselves with the data through in-depth reading to gain a comprehensive understanding; (2) initial codes were generated by breaking the data into meaningful segments to extract relevant points; (3) the codes were organized into categories to identify recurring patterns and potential themes; (4) the emerging themes were reviewed and validated through discussions with expert participants to ensure interpretive accuracy and relevance to the research context; and (5) the researchers defined and named the themes concisely, which were then used to construct the foundational elements of the profiling framework (Baharuddin & Rustan, 2025).

Meanwhile, in the framework analysis, the developed framework was examined and compared using the following dimensions: theoretical foundation and secondary data sources (i.e., journal articles and a selected framework), which were then synthesized to construct a gap analysis matrix. This matrix was used to identify alignments, discrepancies, and development priorities. Ultimately, this entire process would construct a fundamental framework in learning analytics and student profiling.

Trustworthiness and ethical considerations

Member checking was conducted with Universitas Terbuka (UT) stakeholders after preliminary findings were compiled to ensure credibility. Data triangulation was done to enhance dependability. Ethical clearance was sought by informing all informants about the purpose of the research and their rights to confidentiality and withdrawal. Moreover, the researchers utilised OpenAI (ChatGPT) as a writing assistant to help partially edit their text.

Findings

An English education learner profiling framework design

The profiling system for English learners of the EED at Universitas Terbuka (UT) was designed around four major functional phases, namely input, processing, adaptation, and interpretation, each with built-in governance mechanisms to ensure ethical, responsible, and pedagogically aligned implementation. These phases were derived through analysis of existing profiling and learning analytics models and were adapted to suit the practical conditions and institutional infrastructure of Universitas Terbuka (UT).

A series of procedures, including observations and Focus Group Discussions (FGDs), revealed that relevant units manage current data systems at Universitas Terbuka (UT) separately. This conclusion emerged from field observations conducted by the researchers across four key institutional units: *Direktorat Administrasi Akademik dan Kemahasiswaan (DAAK), Pusat Pengelolaan Pembelajaran (PPPb), Pusat Pengujian (Pusjian), and Direktorat Sistem Informasi (DSI).* In addition to these units, further observation was carried out on Universitas Terbuka's Learning Management System (LMS).

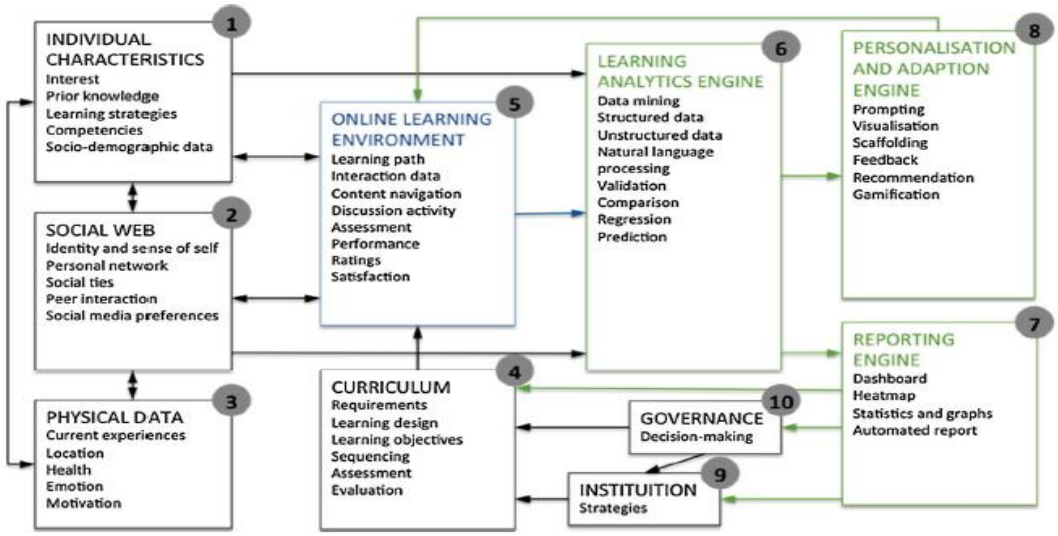
Results from the observations showed that DAAK managed student demographic data, PPPb managed data on learning processes (i.e., LMS behavior), and Pusjian was responsible for overseeing data related to examination results. Each operational unit is accountable for obtaining, recording, managing, and ensuring data availability necessary for further processing by relevant units; datasets are not automatically integrated or processed through a unified profiling engine. Instead, data exchange occurs through a request-based system between units. For example, the DAAK would pull data related to student grades from Pusjian to generate a Student Academic Progress Sheet. Similarly, generating a student's final grade in a course requires scores from the LMS (managed by PPPb) to be combined with the end-of-semester test scores from Pusjian. UG highlighted this data processing system as one of the FGD informants, who stated:

> "Based on my observations, data at UT are managed separately by relevant units, and therefore, a profiling system is yet to be developed, thus personalized learning is still in the making. However, it is important to note that we already have foundation data systems at UT. Our data systems are essentially established, and each unit, such as PPPb, Pusjian, and DAAK can access data according to their respective needs and in line with existing protocols. Having said that, data access is time-bound and subject to specific security measures. Data on student scores obtained during the tutorials (PPPb) and those from the examinations (Pusjian) are eventually combined, translated into final grades and then transferred to the Student Record System (SRS). Overall, this data flow adhered to existing protocols, with clear quality standards and inter-unit data-sharing mechanisms. However, I have not personally encountered or have access to any formal documentation or policy that describes in detail how the system operates, since these would be internal documents."

Three key points must be highlighted above. First, the data system already exists at UT. Second, the data system operates according to the needs of each unit and can be accessed when required. However, data access is subject to timing and depends on when the relevant unit makes the data available (e.g., in the case of LMS and SRS transactional data). In other words, data access remains limited due to concerns around data security. Third, the data system operates based on protocols likely defined individually by each unit. Given this structure, where each unit provides data endpoints either at the beginning or end of the semester, it becomes highly impractical to implement a real-time profiling system during the learning process. This is because the LMS and SRS currently only interact at two points: course registration data are pulled from SRS into the LMS at the start of the semester, and the final scores are then returned to the SRS at the end. The same applies to the integration with Pusjian. At present, there is no continuous or dynamic data exchange between systems. However, an effective profiling system requires integrating various data types—demographic information, LMS behaviour, learning styles, self-regulated learning capacity, and, in the English learners' context, English proficiency—all of which must be captured and processed in real time. This would allow for timely interventions based on the insights displayed through dashboards within the LMS.

When looking into the current dashboard accessible to instructors and students in the LMS, it was necessary to display reports on time, component, event, activity intensity, activity completion, and countries for course management. Other data points recorded in the LMS that can be used to understand learning behaviors were also available, but would need to be manually analyzed to generate actionable interventions. Such interpretive outputs are yet to be presented in the student dashboard meaningfully, which could support personalized learning pathways. Building on the observational findings, the subsequent FGD served as a platform for conducting a comparative discussion centered on a prominent design of a student profiling and Learning Analytics (LA) framework (see Figure 1), specifically the model proposed by Ifenthaler and Widanapathirana (2014).

Figure 1. *Components and relations of the LA framework*



During the FGD, the informant known as IF, who is an educational technology expert, emphasized the following points:

> "The LA framework developed by Ifenthaler has not been seriously contested. The model delineates several key data components, particularly the 'individual characteristics' category, which is further divided into static and dynamic data. These data points are processed through a learning analytics engine and subsequently directed

> to either a personalization and adaptation engine or a reporting engine. When routed to the reporting engine, the data become the basis for targeted interventions. In contrast, data directed only towards the personalization and adaptation engine results in mere informational outputs. Given the current nature of our institutional data, it would be more effective to priorities direct intervention pathways. This has the potential to significantly impact students' learning trajectories through a series of analytic processes. Moreover, considering that the context of this study is English language learning, the inclusion of English proficiency as part of the static data is essential. This addition, encompassing the core language skills, will serve as a distinctive feature of the profiling model we are constructing."

Therefore, the proposed profiling system model incorporates several distinctive data components that are tailored to the context of 21st-century English language learning. One of the key additions included English proficiency as part of the dynamic data, encompassing the four language skills: reading, writing, listening, speaking, and the contextually relevant skill of viewing. In addition, the model integrated variables such as learning style, self-regulated learning, and digital readiness to provide a more holistic representation of the learner. These dynamic components were combined with students' LMS behavior data, ensuring the profiling framework is responds to individual learner characteristics and real-time engagement patterns. As visualized in Figure 2, the framework represented a linear yet interactive progression of how learner data flows through structured processes, resulting in meaningful, tailored educational support.

Phase 1: Input Phase

The first phase answered the question, "What data and from whom is it collected?" This phase involved collecting both static and dynamic data about learners. Static data included demographic and professional information gathered through Universitas Terbuka's Student Registration System (SRS). In contrast, dynamic data would come from ongoing interactions within the learning management system (LMS), diagnostics, and self-assessment tools. At this stage, governance focused on data validation, source reliability, and consent management. Students should be informed about using their data and must provide explicit consent. Data collection instruments would be examined for bias and cultural relevance, per ethical standards and Indonesian regulations.

Phase 2: Processing Phase

The question addressed in this phase was, "How is data transformed into profiles?" During this phase, the profiling engine would compile all the data into learner profiles, updated periodically. A learning analytics engine would perform descriptive and predictive analytics to identify risks, classify learners, and suggest potential needs. Governance in this phase involved auditing algorithm fairness, preventing bias in profiling, and ensuring transparency in decision logic. This process would include routine checks to identify disparities in data representation and model outputs.

Phase 3: Adaptation Phase

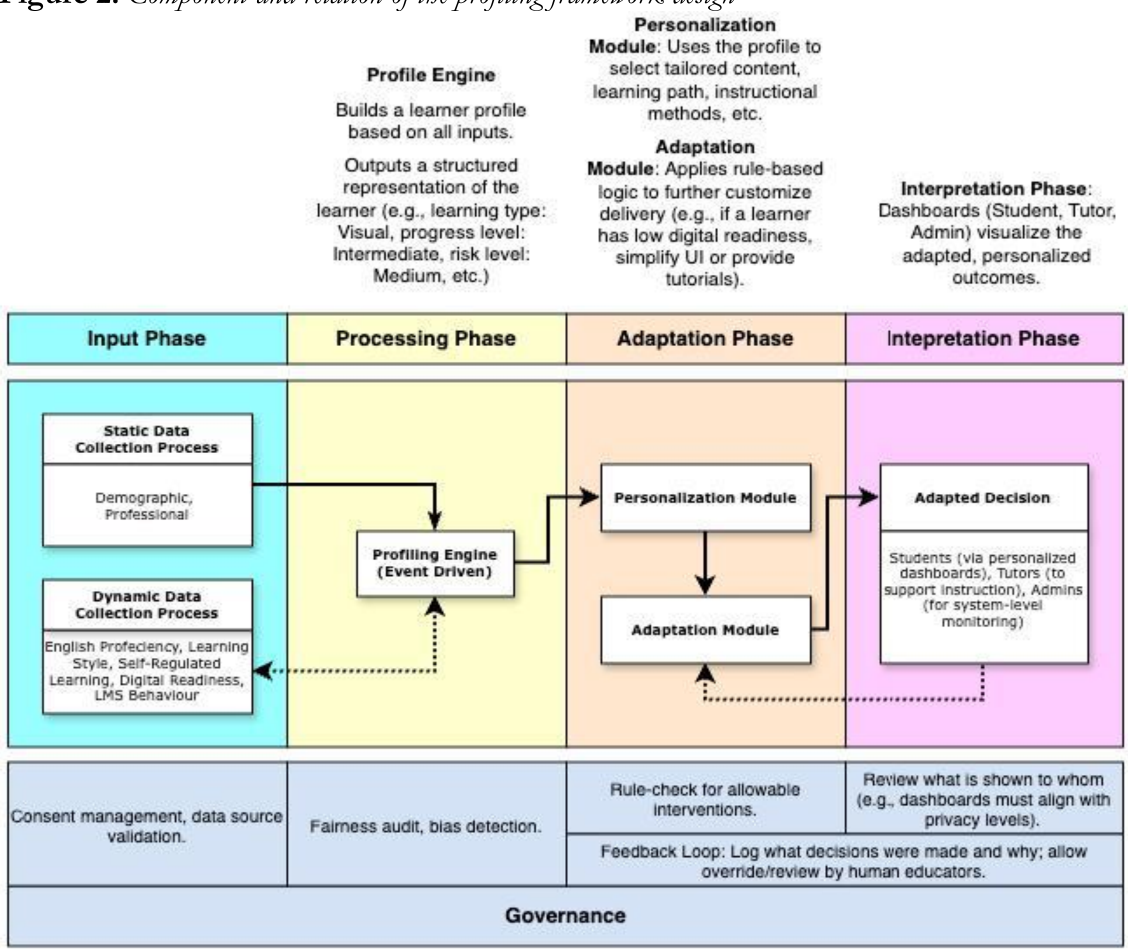
This phase addressed the question, "What recommendations and support are provided?" Based on the learner profiles, the personalization and adaptation modules recommend learning strategies, support services, or content modifications. For example, students with low digital readiness may be given simpler interfaces or digital literacy tutorials. Regarding governance, ethical oversight guarantees the consistent application of rules and provides override mechanisms for human educators to judge inappropriate interventions. A feedback loop logs all automated decisions and justifications.

Phase 4: Interpretation Phase

Lastly, this phase addressed the question, "How do stakeholders use the data?" Dashboards and visualizations translate profiling outcomes into actionable insights for students, tutors, and administrators. For example, tutors may view class-level analytics, while students receive personalized reports. The governance role in this phase emphasizes privacy management and data visibility rights, ensuring each user sees only the data appropriate to their role. Dashboard access is role-specific and aligned with Universitas Terbuka's data protection policy.

Across all phases, the governance layer functions as a critical safeguard, ensuring data usage remains transparent, secure, equitable, and pedagogically aligned. Its integration reinforces accountability, protects learner agency, and upholds institutional integrity.

Figure 2. *Component and relation of the profiling framework design*



Institutional mapping of current infrastructure and readiness

To ensure that the proposed profiling framework is grounded in institutional reality, a mapping exercise was conducted to assess the availability, accessibility, and readiness of relevant data components at Universitas Terbuka (UT). This mapping process was conducted collaboratively with the FGD participants, who were provided with relevant materials and guided through a structured discussion. As a result, several key themes emerged from the analysis, forming the basis for the conceptual mapping presented in this research (see Table 1).

Moreover, the mapping cross-referenced each significant component of the framework—across all four functional phases—with Universitas Terbuka's existing systems, including the Student Registration System (SRS), Moodle-based Learning Management System (LMS), and MyUT portals for students and lecturers. While some data (e.g., demographics, academic history) are already routinely collected, other dimensions, such as psychological readiness and learning style profiling, require additional instruments or system upgrades. This exercise revealed both strengths and areas for development, as summarized in Table 1.

Table 1. *Mapping current infrastructure and data availability to the multidimensional model*

Framework Component	Current Availability at UT	Gap / Recommendation
Demographics	Captured via SRS (Student Registration System)	None
Academic History	Accessible through academic archives and LMS	Integration with the profiling system needed
Professional Experience	Collected during registration	Add structured fields for teaching context and experience
Digital Literacy	Available by demand	Include a digital readiness checklist or self-report tool
English Proficiency	CEFR-based not fully implemented	Link to placement tools or a short diagnostic test
Learning Styles & Modality	Currently captured during orientation sessions	Deploy a validated learning style inventory during onboarding
Psychological Readiness	No data	Add self-regulated learning readiness scale (e.g., SRLQ)
LMS Interaction Data	Moodle logs available	Need an analytics pipeline to interpret data meaningfully
Profiling Engine	Not yet developed	Build an integrated profiling database and input portal
Analytics Engine	Limited use in dashboards	Partner with the IT team to develop statistical and ML models

Framework Component	Current Availability at UT	Gap / Recommendation
Personalization Engine	In progress	Rule-based recommendations can be implemented first, ML later
Student Dashboards	MyUT-Student	Develop tailored visualizations linked to profiles
Lecturer/Program Dashboards	MyUT-Dosen	Design role-based views for tutoring support and intervention planning
Data Governance & Consent	Ethics protocol exists, but not student-facing	Add consent at login and training for ethical use
Stakeholder Literacy Training	Already in place	Workshops and guides for tutors, students, and program heads

The mapping revealed that Universitas Terbuka (UT) already has a strong infrastructure to support basic profiling, especially regarding student identity and academic records. However, to realize the full potential of the profiling framework, targeted enhancements such as integrating validated psychological scales, developing a rule-based personalization engine, and visual analytics dashboards should be developed.

Discussion

The findings from this research have illustrated how a carefully designed profiling system can meet the multifaceted needs of open and distance learners, particularly within a large-scale institution like Universitas Terbuka (UT). Grounded in theory and contextual realities, the proposed framework integrates key learner dimensions and system components in a phased model that supports continuous learning improvement and institutional decision-making.

By focusing on core learner attributes, i.e., digital readiness, English proficiency, learning styles, and readiness for independent learning, the framework addressed foundational elements critical to academic success in an online environment. These dimensions were not selected arbitrarily but through a rigorous synthesis of contemporary profiling literature and field-based requirements. This aligned with Adnan and Anwar (2020), who emphasized the role of learner autonomy and technological preparedness in online education success. Similarly, incorporating digital readiness is consistent with recent studies (Fernandez et al., 2022; Kim et al., 2019). They underscored the importance of digital competence (i.e., skill, attitudes, and knowledge) in e-learning environments.

The functional architecture of the system, comprising the Input, Processing, Adaptation, and Interpretation phases, was a direct response to the call for transparency, adaptability, and ethical awareness in student profiling (Al-Zahrani & Alasmari, 2023; Wise et al., 2021). It was built upon the theoretical constructs offered by Ifenthaler and Widanapathirana (2014), who emphasized the integration of profiling, analytics, and personalization engines. Furthermore, the model adopted ethical safeguards highlighted by Greller and Drachsler (2012) and reaffirmed stakeholder-driven analytics frameworks suggested by Chatti et al. (2012), integrating them into the governance layer.

Rather than treating governance as a peripheral or reactive element, this framework situated it at the core of every phase in the profiling process—from consent protocols in the Input phase to access control in the Interpretation phase. This was aligned with growing scholarship emphasizing that fairness, equity, and transparency must be embedded in system design. According to Uttamchandani and Quick (2022), even well-intentioned analytics tools can perpetuate biases unless they are critically assessed for equity and include mechanisms to monitor and mitigate algorithmic or human bias.

Recent research underscored that learning analytics systems often fail to produce fair outcomes for historically underserved populations unless they incorporate explicitly designed fairness protocols (Prinsloo et al., 2024). These insights affirmed why our profiling framework includes role-based access controls, consent management, bias audits, and transparency reports. A systematic review by Uttamchandani and Quick (2022) warned that data-driven decision-support tools may reinforce inequities unless fairness criteria are proactively integrated into algorithmic design. Echoing this, our approach embeds governance into every phase to ensure ethical reflection. This functionality was built into data protocols, analytics logic, and the stakeholder dashboard and was not added later.

The framework has illustrated that effective learner profiling in ODL contexts requires more than data collection; it demands a coherent, ethically informed system architecture that translates insights into targeted support and policy improvements. This research has contributed a scalable, context-sensitive model for learner profiling that reflected best practices in recent literature and could guide future implementation and refinement within Universitas Terbuka (UT) and similar institutions.

Conclusions and Recommendations

This research outlines the conceptual foundation and institutional alignment for a multidimensional profiling framework to enhance personalization in English language learning at Universitas Terbuka (UT). The proposed model was carefully designed based on a synthesis of relevant learning analytics frameworks and the specific needs of a large-scale open and distance education institution. The framework integrated various learner dimensions, processing mechanisms, interpretive tools, and ethical considerations to support adaptive learning, aligning with Universitas Terbuka's operational capacity and learner diversity.

The findings indicated that while Universitas Terbuka (UT) already possesses strong foundational systems for academic and demographic profiling, there remain critical gaps in profiling tools related to learner behavior, readiness for independent learning, and digital proficiency. To move forward, Universitas Terbuka (UT) should prioritize developing and validating diagnostic instruments to measure these aspects. In tandem, it would be necessary to integrate analytics engines and interactive dashboards into the existing learning management system to use these data points in a timely and pedagogically meaningful way.

Universitas Terbuka (UT) should look into adopting a phased approach to implementation, beginning with pilot testing the profiling tools in selected English language courses. Staff and tutors must be trained to interpret profile data and ethical and pedagogical practices associated with its use. Data governance protocols must also be formalized, including consent procedures and protection mechanisms. In doing so, Universitas Terbuka (UT) can set a precedent for scalable, ethical, and learner-centered profiling systems that are responsive to the needs of distance learners in Indonesia and potentially similar educational contexts.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest.

Acknowledgments

The authors gratefully acknowledge Universitas Terbuka (UT) and its English Education Department for supporting the infrastructure review and institutional interviews necessary for this research.

References

Adnan, M., & Anwar, K. (2020). Online learning amid the COVID-19 pandemic: Students perspectives. *Journal of Pedagogical Sociology and Psychology*, *2*(1), 45-51. https://doi.org/10.33902/jpsp.2020261309

Al-Zahrani, A. M., & Alasmari, T. (2023). Learning analytics for data-driven decision making: enhancing instructional personalization and student engagement in online higher education. *International Journal of Online Pedagogy and Course Design*, *13*(1), 1-18. https://doi.org/10.4018/IJOPCD.331751

Baharuddin, A. F., & Rustan, D. R. H. P. (2025). Investigating intrinsic and extrinsic factors influencing English self-learning motivation among Indonesian university students. *Jo-ELT (Journal of English Language Teaching) Fakultas Pendidikan Bahasa & Seni Prodi Pendidikan Bahasa Inggris IKIP*, *12*(1), 103-116. https://doi.org/10.33394/jo-elt.v12i1.15386

Braun, V., & Clarke, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The Psychologist*, *26*(2), 120-123.

Camarillas, N. G. (2018). Perceptual learning styles and multimodality in EFL education: the digital learning style and social networks. *Fòrum de Recerca*, (23), 311-325. https://doi.org/10.6035/forumrecerca.2018.23.21

Carow, S., Schmitz, P., & Pretorius, R. (2023). Spatial perspectives on student profiling to inform open distance e-learning (ODeL) in various geographical contexts: a case study from the Global South. *Discover Sustainability*, *4*(1), 28-35. https://doi.org/10.1007/s43621-023-00143-9

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, *4*(5–6), 318-331. https://doi.org/10.1504/IJTEL.2012.051815

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed method approaches*. California: SAGE Publications.

Du Plooy, E., Casteleijn, D., & Franzsen, D. (2024). *Personalized adaptive learning in higher education: A scoping review of key characteristics and impact on academic performance and engagement*. *Heliyon*. https://doi.org/10.1016/j.heliyon.2024.e39630

Fernandez, A. I., Al Radaideh, A., Sisodia, G. S., Mathew, A., & del Río, J. A. J. (2022). Managing university e-learning environments and academic achievement in the United Arab Emirates: An instructor and student perspective. *PLoS ONE*, *17*(5), 1-11. https://doi.org/10.1371/journal.pone.0268338

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology and Society*, *15*(3), 42-57.

Guo, S., Yao, H., & Lee, J. S. (2025). A latent profile analysis of informal digital learning of English: Emotion regulation strategies as predictors and willingness to communicate as outcomes. *System*, *132*(2), 21-27. https://doi.org/10.1016/j.system.2025.103686

Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning*, *19*(1–2), 221-240. https://doi.org/10.1007/s10758-014-9226-4

Ifenthaler, D., & Yau, J. Y. K. (2020). Utilising learning analytics to support study success in higher education: a systematic review. *Educational Technology Research and Development*, *68*(4), 1961-1990. https://doi.org/10.1007/s11423-020-09788-z

Isman, S. M., Nugraha, B., Hadianti, S., Zuhairi, A., & Rasyki, A. (2024). Profiling student based on academic performance: identifiying key characteristics in an online learning environment. *International Conference on Innovation in Open and Distance Learning, 5,* 1233-1249.

Kim, H. J., Hong, A. J., & Song, H. D. (2019a). The roles of academic engagement and digital readiness in students' achievements in university e-learning environments. *International Journal of Educational Technology in Higher Education*, *16*(1), 1-18. https://doi.org/10.1186/s41239-019-0152-3

Kuruppuarachchi, K. A. J. M., & Karunanayake, K. O. L. C. (2017). Student characteristics and knowledge on ODL concepts at first registration: A case study from OUSL. *Asian Association of Open Universities Journal*, *12*(1), 41-51. https://doi.org/10.1108/AAOUJ-11-2016-0004

Latif, L. A., Subramaniam, T. T., & Khatab, Z. A. (2020). *Learner profiling towards improving learner success*. Singapore: Springer Nature Singapore. https://doi.org/10.1007/978-981-15-6591-5_14

Lin, L. C., & Tsou, W. (2025). Technology-mediated online EMI professional development: Developing faculty self-efficacy and teaching practice through the community of inquiry framework. *System*, *133,* 10-27. https://doi.org/10.1016/j.system.2025.103727

Ma, L., & She, L. (2024). Self-regulated learning and academic success in online college learning. *Asia-Pacific Education Researcher*, *33*(3), 519-533. https://doi.org/10.1007/s40299-023-00748-8

Ozturk, A., Kumtepe, A. T., & Aydın, S. (2019). *Mapping the learner profiles in open and distance learning.* Madrid: UNED.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research*, *31*(6), 459-470. https://doi.org/10.1016/S0883-0355(99)00015-4

Prinsloo, P., Khalil, M., & Slade, S. (2024). Learning analytics as data ecology: a tentative proposal. *Journal of Computing in Higher Education*, *36*(1), 154-182. https://doi.org/10.1007/s12528-023-09355-4

Siemens, G., & Long, P. (2011). Penetrating the Fog: Analytics in learning and education. *EDUCAUSE Review*, *46*(5), 1-11.

Tsai, C. L., Cho, M. H., Marra, R., & Shen, D. (2020). The self-efficacy questionnaire for online learning (SeQoL). *Distance Education*, *41*(4), 471-489. https://doi.org/10.1080/01587919.2020.1821604

Uttamchandani, S., & Quick, J. (2022). An introduction to fairness, absence of bias, and equity in learning analytics. *SOLAR, 10*(4), 205-212.

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*. *89*, 98-110. https://doi.org/10.1016/j.chb.2018.07.027

Wise, A. F., & Vytasek, J. (2017). *Learning analytics implementation design*. Canada: Simon Fraser University. https://doi.org/10.18608/hla17.013

Wise, A. F., Knight, S., & Ochoa, X. (2021). What makes learning analytics research matter. *Journal of Learning Analytics*, *8*(3), 1-9. https://doi.org/10.18608/jla.2021.7647

Biographical Notes

SUCI MADIARTI ISMAN is a lecturer at Universitas Terbuka, Banten, Indonesia.

IZZUL FATAWI is a lecturer at Universitas Terbuka, Banten, Indonesia.

DWI REZKI HARDIANTO PUTRA RUSTAN is a lecturer at Universitas Terbuka, Banten, Indonesia