FAHIM: Flexible, Insight-Driven, Gamified Assessment Hub for Higher Education

*[1]Ahmad Fahimi Amir,

[1]Universiti Pertahanan Nasional Malaysia

*Corresponding Author

DOI: https://doi.org/10.47772/IJRISS.2025.925ILEIID000019

Received: 23 September 2025; Accepted: 30 September 2025; Published: 04 November 2025

ABSTRACT

FAHIM is a web-based platform application that helps higher education instructors design fair, outcomes-aligned assessment in minutes. Its core innovation is the tight coupling of a reusable Question Bank tagged by subject, question type, difficulty, Bloom's level, vocabulary/CEFR tier, and marks with an intuitive drag-and-drop Paper Builder and instant analytics. The app also features an in-app guidance to support users in developing items, drafting question stems and options, and a scoreboard with badges gamifies contribution by tracking progress against peers and celebrating milestones. The app reduces preparation time and provides better balance across difficulty and cognitive levels, clearer alignment to outcomes, and greater transparency for moderation. The gamified scoreboard drives engagement and continuous improvement, while AI assistance lowers the learning curve. Together, FAHIM turns exam design into repeatable, data-informed practice rather than a last-minute craft and demonstrates how metadata, lightweight analytics, and an embedded AI assistant can advance accuracy, clarity, and consistency in higher education assessment.

Keywords: question bank, assessment design, educational analytics, quality assurance, higher education

INTRODUCTION

In higher education, the assessment process is integral to educational quality, demanding a robust approach to ensure validity, reliability, and fairness. These core principles underpin the efficacy and credibility of assessments. Validity refers to the extent to which an assessment accurately measures what it is intended to measure, ensuring alignment with learning outcomes and objectives (Flores et al., 2014). This involves using clear and relevant criteria, such as rubrics that delineate specific expectations and standards for student performance, providing a clear roadmap for both educators and learners (Ragupathi & Lee, 2020). Reliability is concerned with the consistency of assessment results across various contexts and different groups of students. This principle ensures that an assessment produces similar outcomes under consistent conditions and techniques (O'Neill, 2017).

Cognitive rigor, as exemplified by Bloom's taxonomy, plays a vital role in promoting higher order thinking by encouraging educational activities that move beyond basic recall of information to more complex processes such as application, analysis, synthesis, and evaluation. Bloom's taxonomy provides a structured framework that helps educators design curriculum and assessments with clear learning objectives that foster critical thinking and problem-solving skills. This taxonomy is particularly important for developing students' abilities to apply knowledge in new contexts and to engage in reflective learning, which are essential components of higher-order thinking (Bissell & Lemons, 2006; Zaidi et al., 2018).

In relation to language assessment, the Common European Framework of Reference for Languages (CEFR) serves as a critical benchmark. It provides a standardised framework that outlines language proficiency levels from A1 to C2, thereby facilitating the measurement and comparison of language skills across different contexts. The CEFR sets clear standards for what learners are expected to attain at each proficiency level, promoting consistency and transparency in language assessment and curricula. This helps educators and testing agencies design assessments that accurately reflect a learner's language abilities, ensuring a comprehensive evaluation of linguistic skills (Gaillat et al., 2021; Wisniewski, 2017).

Problem Statement

Poor assessment design significantly affects both student learning outcomes and equity. Traditional assessment methods have often been criticized for their inability to accommodate the diverse learning needs, values, and preferences of students, leading to disengagement and poor learning outcomes (Smith et al., 2024). Inappropriate assessment tasks that do not align with contemporary learning theories may further exacerbate these issues, failing to promote equity despite the intentions to do so (Kang & Furtak, 2021).

To make matter worse, instructors often resort to ad hoc or last-minute assessment design due to various systemic and practical challenges. One primary issue is the lack of sufficient time and resources for thorough planning and development of assessments. Educators frequently face time constraints due to their numerous responsibilities, including class preparation, student support, and administrative duties, leaving little opportunity for careful assessment design (Cheng et al., 2004; Kerneža & Zemljak, 2023).

Educators under pressure often have to produce assessments quickly, which might not thoroughly align with the intended learning goals. The absence of appropriate digital tools and technologies to aid in assessment design exacerbates this issue, as these tools can facilitate better alignment through automation and standardisation of content (Akintayo et al., 2024). This misalignment can undermine the validity and reliability of the assessment outcomes, skewing data on student performance and impairing the ability to make informed decisions on student progress and curriculum effectiveness (Brown, 2022).

Furthermore, educators may lack adequate support systems and collaborative frameworks that encourage sharing best practices and innovative approaches in assessment design. Structural deficiencies such as limited collaboration opportunities can prevent teachers from leveraging peer insights and experiences to enhance their assessment strategies (Perry et al., 1999), with increasing challenges in digital assessment (Viberg et al., 2024).

Constructive alignment, a framework that aligns learning activities and assessments with intended learning outcomes requires that the assessments directly measure the skills or knowledge described by the learning outcomes, thus ensuring clarity and purpose in educational tasks (Ali, 2018; Biggs & Tang, 2014). Strategically aligning learning outcomes with assessments is directly related to Sustainable Development Goal 4 (SDG 4), by fostering educational environments where assessments are used to truly reflect and achieve intended educational aims, thereby enhancing educational quality and providing all students with the opportunity to succeed (Volkov, 2023).

Language demands, particularly as defined by frameworks like the Common European Framework of Reference for Languages (CEFR), significantly influence the accessibility and fairness of assessments by standardising language proficiency descriptions, enabling assessments to be aligned to these levels. For instance, the CEFR provides a guideline for designing language tests that assess speaking, listening, reading, and writing skills, which can contribute to fairness by maintaining consistency across various test formats (Baharum et al., 2021).

In Malaysia, educators were found to lack knowledge on how to integrate CEFR into teaching practices (Uri, 2025). Issues were raised on the need to normalise academic practice to improve the CEFR-aligned educational program (Lee et al., 2022).

Thus, systemic improvements, such as better tools for assessment design and more comprehensive training for educators, can mitigate these issues and foster more accurate and meaningful alignment between learning outcomes and assessments. With these goals in mind, the aim is to develop an application with the following objectives:

Objectives

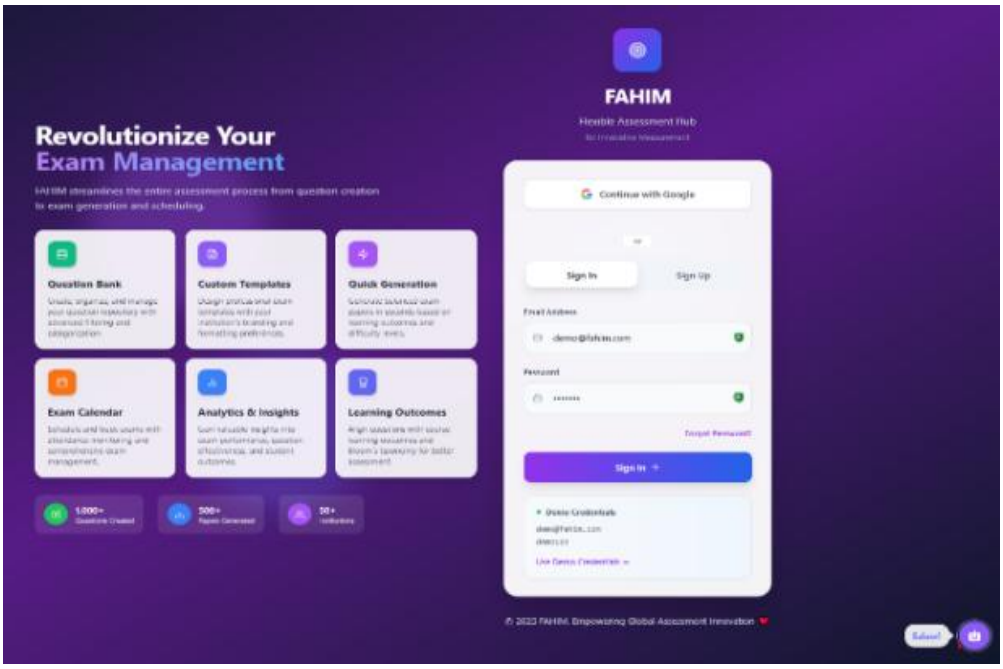
1. To enhance the alignment between learning outcomes and assessment items in higher education by leveraging a metadata-rich, reusable question bank tagged with Bloom's cognitive levels, difficulty, and CEFR-based vocabulary tiers, thereby supporting educators in designing valid, transparent, and inclusive assessments.

2. To improve the efficiency, consistency, and pedagogical quality of assessment design through an integrated web-based platform that combines intuitive drag-and-drop authoring, AI-assisted item development, real-time analytics, and gamified engagement mechanisms.

PRODUCT DESCRIPTION & METHODOLOGY

The development of this application was based on a conceptual framework (later validated and named SEER Framework) developed by the author, focuses on having a supportive, effective, ethical, and relevant in order to provide a clear, operational vocabulary to guide design decisions from ideation to implementation.

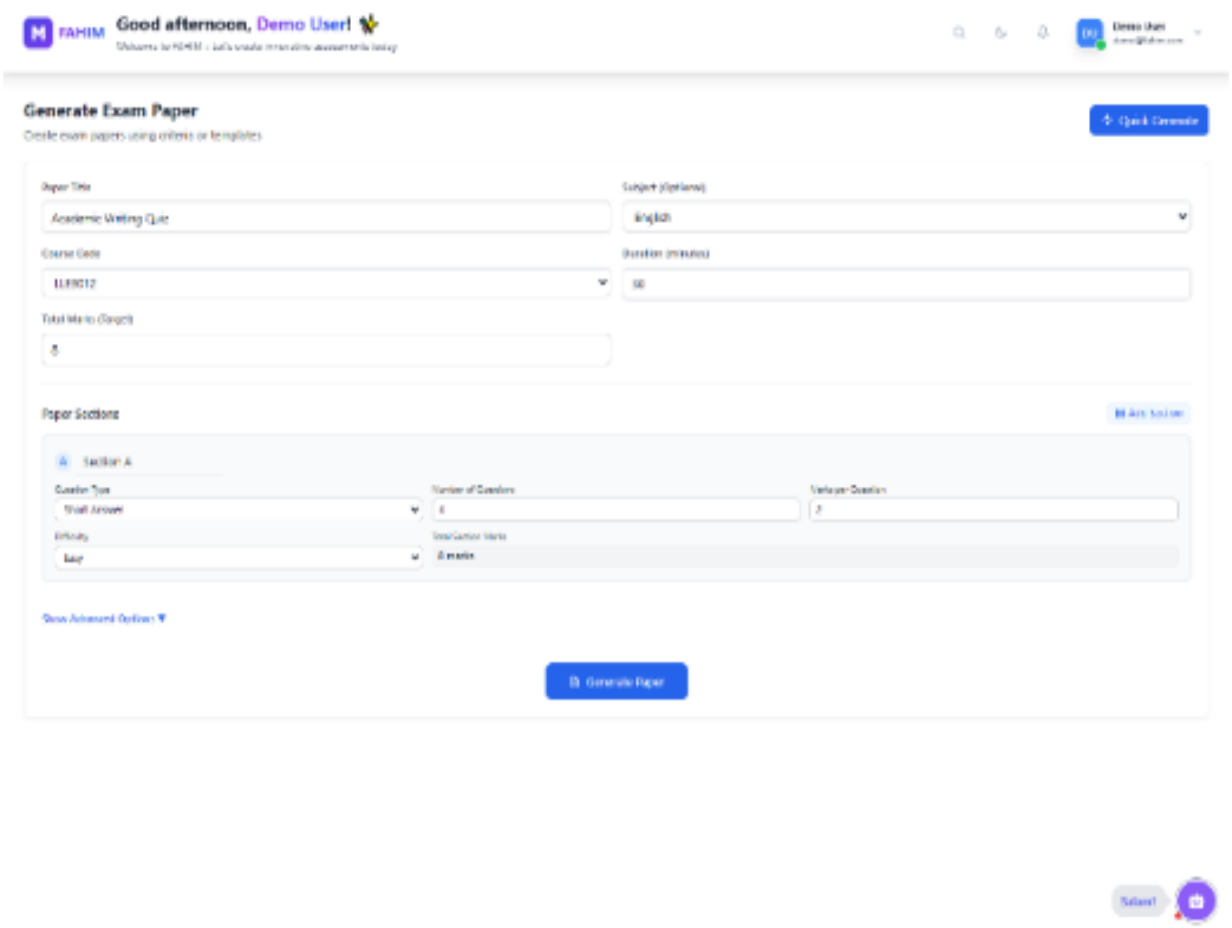
Figure 1: S.E.E.R Framework (Amir et al., 2024)



Other documents that were also referenced are frameworks and benchmarking by agencies that are responsible in providing quality assurance particularly in higher education such as Malaysian Qualification Agency's (MQA) Guidelines to Good Practices and Council of Europe's CEFR descriptors. These references were used to provide guidelines for the app's automatic mapping of question items to Bloom's taxonomy and CEFR tiers.

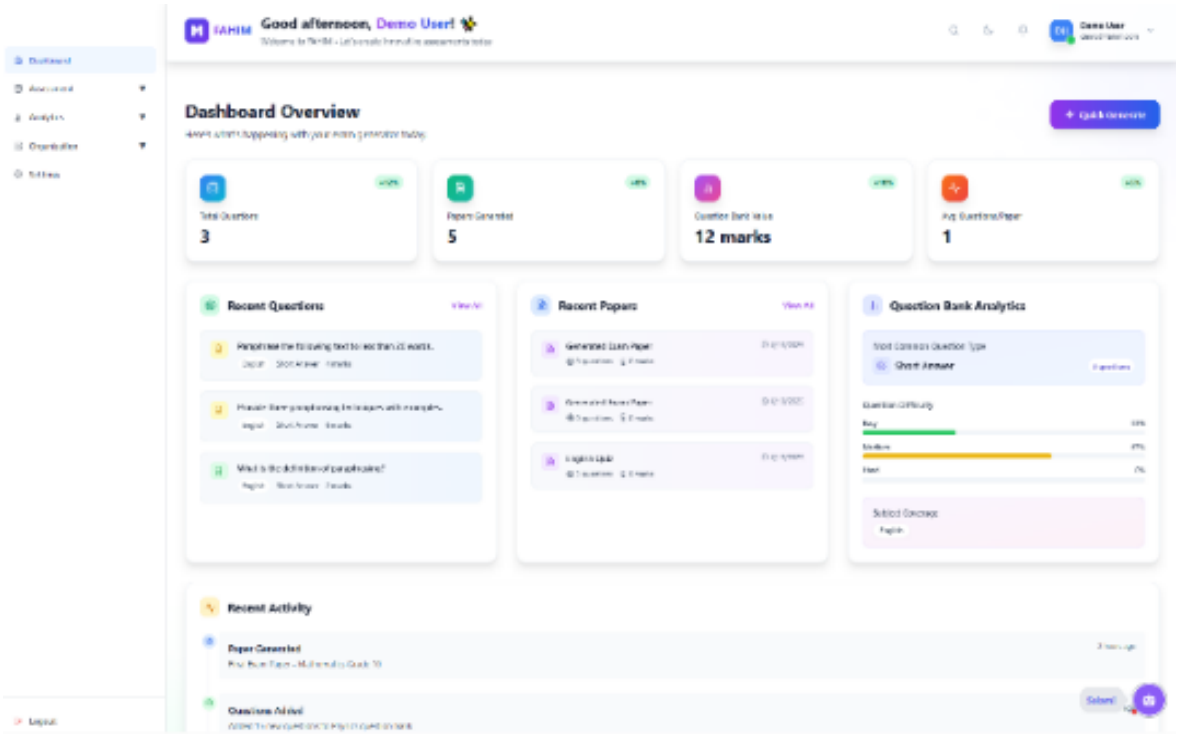
The application aptly named FAHIM (which means to comprehend or to understand in Arabic) was built in React 18 + Vite environment, written in JavaScript and edited through open-source Visual Studio Code (VS Code) with Supabase as the backend and GitHub as the repository for the source code.

Figure 2: FAHIM's login page



The login page explains all the available features of the app. Registered users can access the app for the purpose of storing and analysing of question items and will only be accessible by the users who created an account. The data is kept safe via HTTPS encryption.

Figure 3: FAHIM's dashboard for overview

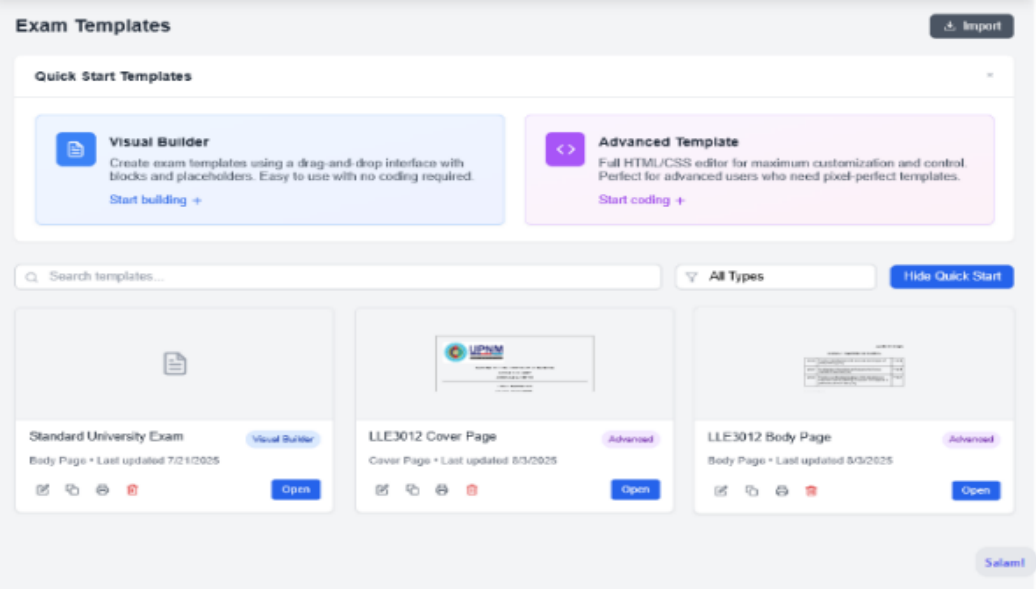


The dashboard allows for a quick snapshot on the recent activities and progress that a user has made on the question bank database. Data such as number of items, papers generated and total marks in the question bank are automatically calculated and are displayed in the dashboard.

Figure 4: Quick generator for rapid exam paper generation

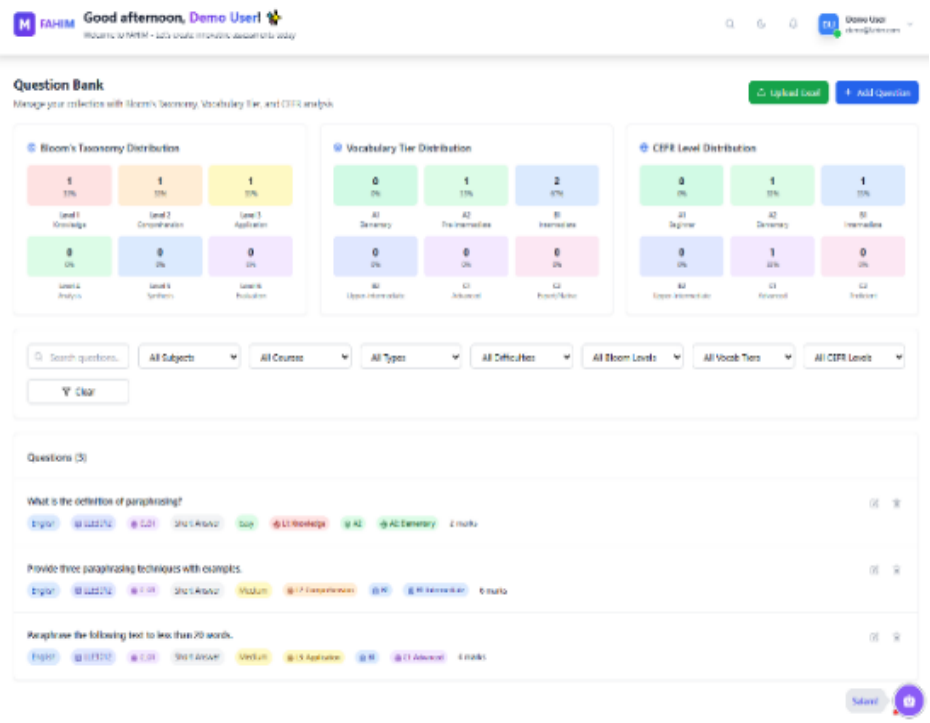
A user can quickly generate an exam or a quiz paper using a feature called 'Quick Generate'. FAHIM will pull available resources from the question bank and generate the paper based on user's settings. This process will take less than 10 seconds.

Figure 5: Generate paper using advanced templates



If users are particular about the formatting of the paper, they can opt for advanced templates by using either the visual builder or HTML builder. HTML builder provides the most precise formatting options but requires user to have a bit of knowledge in HTML codes.

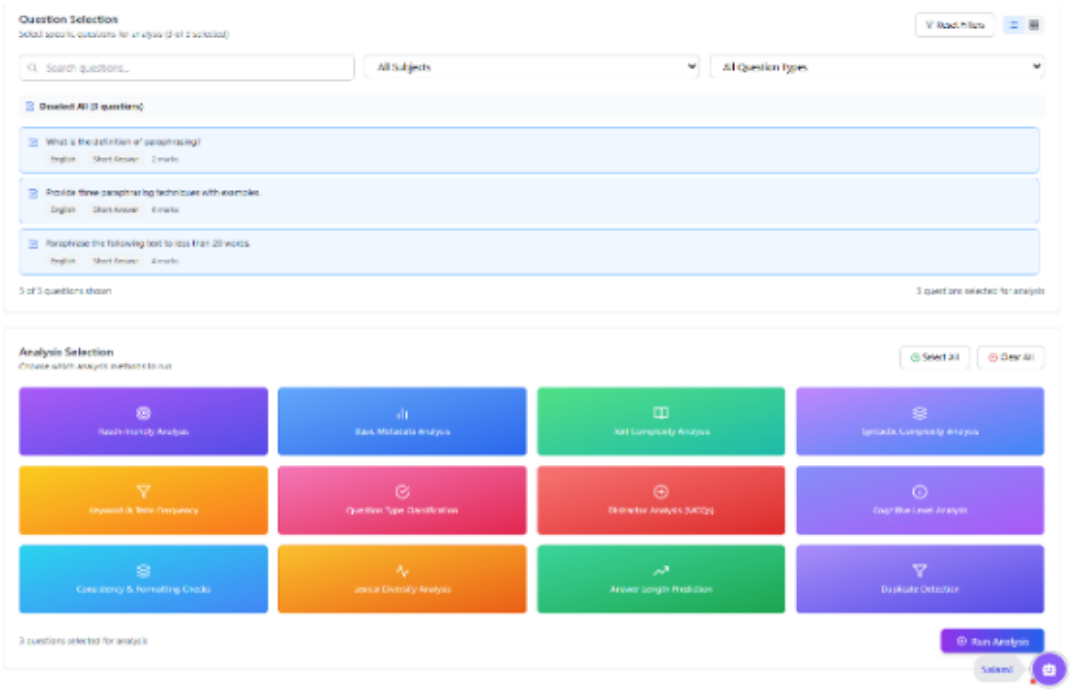
Figure 6: The question bank dashboard



The question bank page allows user to see the total number of items in their collection which are categorised based on Bloom's taxonomy and CEFR level. Systematic tagging also allows users to identify each item's subject, course code, course learning outcome, question type, difficulty, Bloom's taxonomy level, CEFR level and assigned marks (Figure 6).

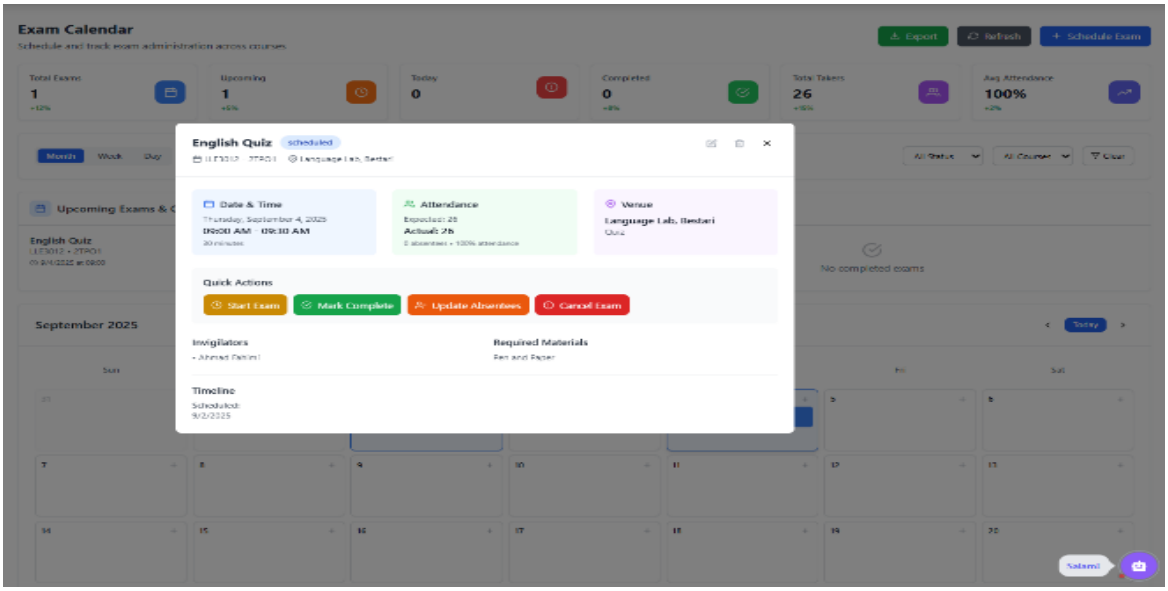
To add items into the database, users can either add manually or uploading it via excel file. If adding items via excel file, uploading 100 items will take less than 15 seconds. User can easily follow the provided template and replace the columns with the correct data.

Figure 7: Item analysis



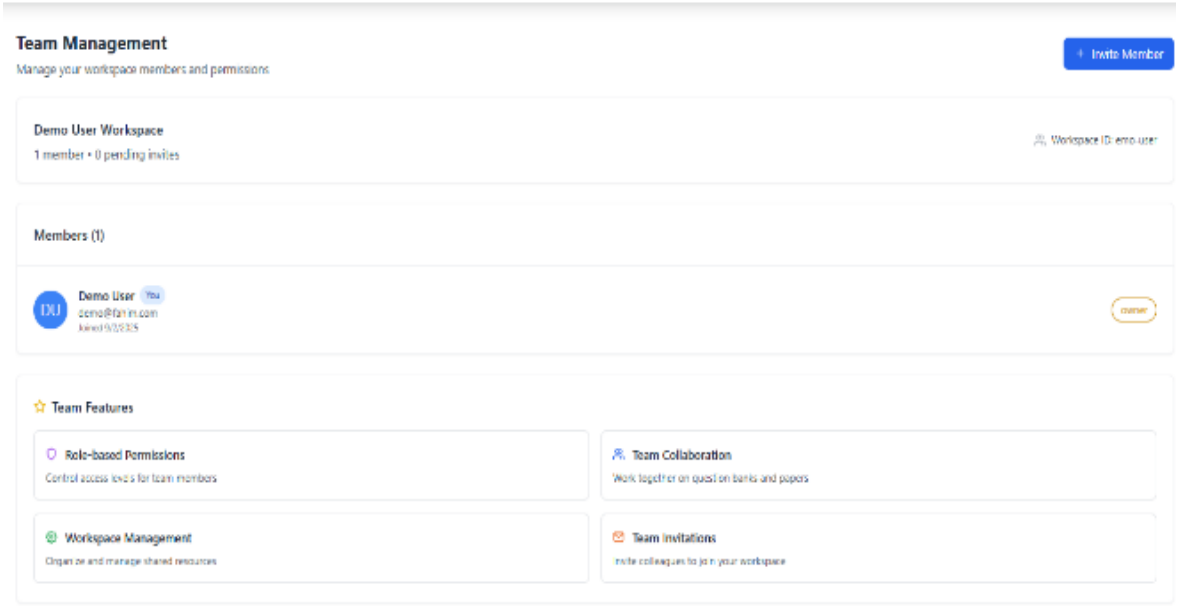
Users have the ability to run an analysis test on the question bank to get a clear picture of items in their collection. These types of analysis help users to manage their question bank by allowing users to get a quick snapshot of the items in their collection, assess readability of the questions, identify potential redundancy or near-duplicate items, maintain healthy item-type mix, find typos or formatting mistakes and accurately maps to Bloom's levels via verb cues.

Figure 8: Exam calendar to better manage exam schedules



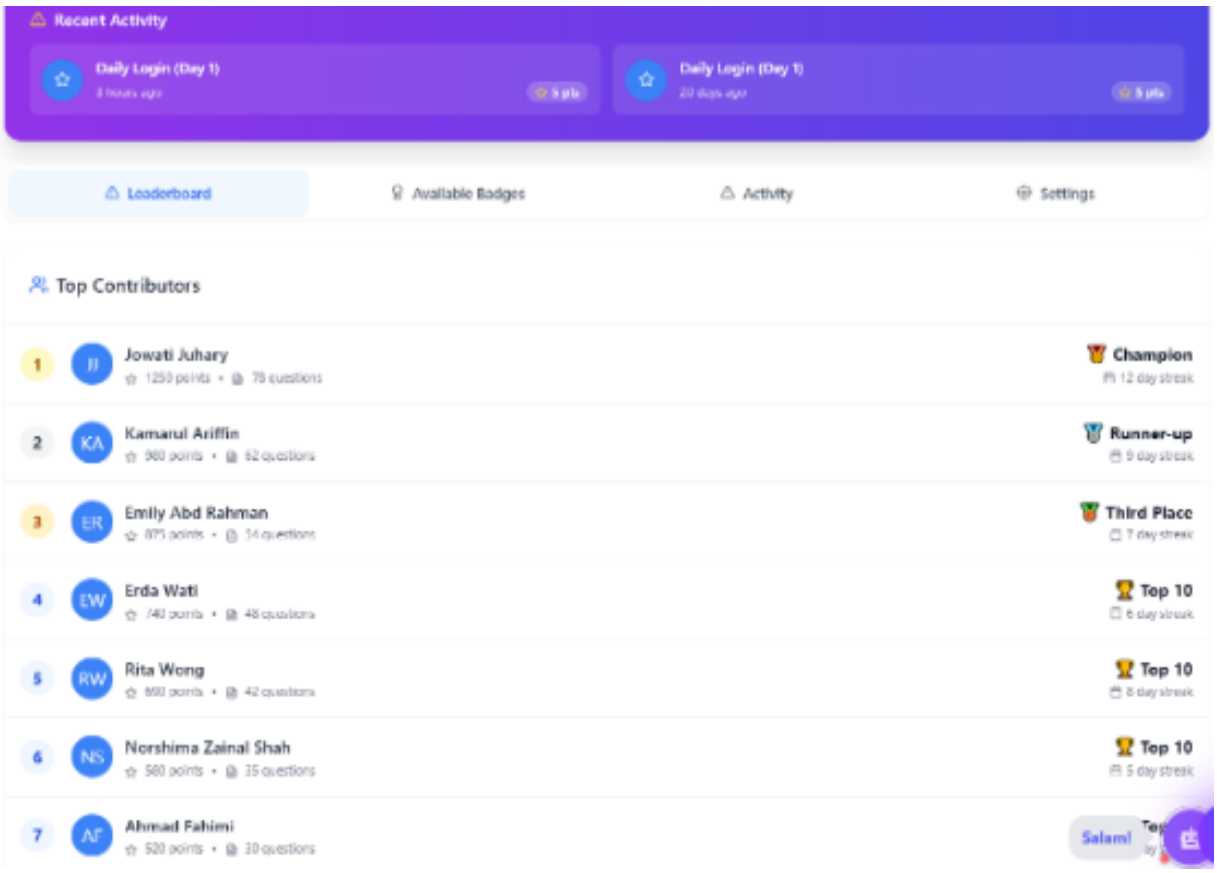
FAHIM provides a way for users to manage their exam schedule via an Exam calendar. Users simply pick a location, date and time, and provide the number of test takers. A bell notification will be shown at the top of the page when an upcoming exam is near and once the exam has ended, user can update the list of absentees for record purposes.

Figure 9: Team management for collaboration



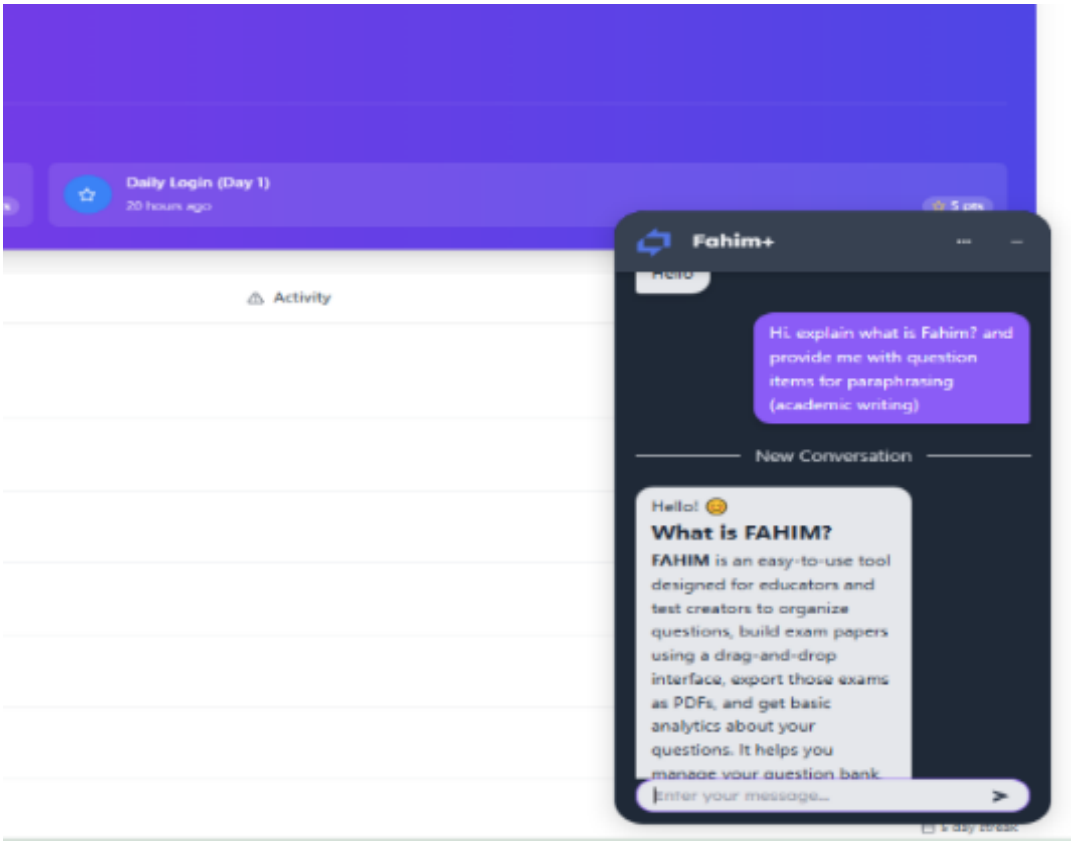
FAHIM allows users to create a team in their workspace so that they can work together in increasing the number of items in their collection. This collaboration allows for better management of items and improves items validity.

Figure 10: Leaderboard for healthy competition



FAHIM gamifies the hectic process of managing items by rewarding users who put more effort in expanding their question bank. By achieving a variety of badges and frequenting the application, users can gain more points which help them to be listed in the top contributor's leaderboard. This form of recognition provides the extra incentives for educators to increase the number for items in the question bank.

Figure 11: Fahim+ AI chatbot for assistance



A specialised AI chatbot called Fahim+ is also available to users when navigating the application. The chatbot was trained with relevant documents and will answer any enquiries regarding FAHIM application and related tasks.

The following table summarises the features of the app and the applied concepts:

Table 1 Features & Concept

Features	Linked Concept	Descriptions
Tagged Question Bank (Bloom's, CEFR, Learning Outcome)	Taxonomic alignment (Anderson et al., 2000) CEFR (Prajapati, 2022) Metadata (Mathew, 2020)	-facilitates the classification of questions according to cognitive complexity -provides a comprehensive framework for understanding language proficiency across different contexts - essential for organising and retrieving detailed information about each question
Drag-and-drop Paper Builder + Templates	Cognitive load theory (Sweller, 1994 Technology Acceptance Model 3 (TAM3) (Venkatesh & Bala, 2008)	- support the design of user-friendly interfaces by minimizing extraneous load through intuitive drag-and-drop functionalities - supports user acceptance and sustained engagement by reducing perceived effort
Instant Dashboard Analytics & Moderation Support	Educational data mining (Romero & Ventura, 2020)	-leverages educational data mining techniques to provide real-time insights into the question bank - assists in maintaining assessment quality by identifying anomalies

AI Chatbot for Onboarding & Item Refinement, Dashboard	AIED (Holmes & Tuomi, 2022)	- personalised guidance and facilitates item refinement through adaptive feedback
Gamified Scoreboard & Badges	Self-determination theory (Ryan & Deci, 2000)Gameful design (Deterding, 2015) Effectiveness of badges and leaderboards (Balci et al., 2022; Luo, 2024)	- enhances motivation by fostering intrinsic engagement and promoting sustained participation
Calendar & Assessment Tracking	Time management & planning in teaching (Szwarc et al., 2018)	- facilitate efficient scheduling and timely monitoring of assessments
Spreadsheet Import & PDF Export Item Bank	Interoperability Item bank (Sclater & MacDonald, 2004)	- enable seamless integration with existing systems, facilitating efficient data exchange

Comparative Analysis With Existing Platforms

FAHIM was benchmarked against question bank software developed by world-leading companies for a comparative analysis. Comparison against these gold standards will highlight the novelty that FAHIM brings into the market. Speedwell, Maxexam, Synap, Moodle and Examarius were selected for this purpose, and their respective features advertised in their websites are compared as shown in Table 2 below.

Table 2 Platform Feature Comparison

Platform	**Question Bank Features**	**Assessment Creation**	**Analytics & Insights**	**Collaboration Features**	**Gamification**	**AI Integration**
FAHIM	Bloom's taxonomy, CEFR levels, Learning outcomes, Difficulty tagging	Drag-and-drop Paper Builder, Quick Generator, Templates	Real-time dashboard, Item analysis, Moderation support	Team management, Gamified leaderboard, Peer tracking	Badges, Leaderboards, Point system, Achievement tracking	AI chatbot (Fahim+), Item refinement, Onboarding assistance
Examarius	Hierarchical organization, Bulk editing, Cross-search filtering	TestBuilder with drag-and-drop, Intuitive interface	Question performance statistics, Usage tracking	Multi-user access, Collaborative editing	Limited gamification elements	None mentioned
Speedwell	Statistical analysis, Performance tracking, Blueprinting	Advanced blueprinting, Curriculum alignment	Statistical analysis, Performance identification	Remote authoring, Access controls, Permissions	None mentioned	None mentioned
Synap	Custom question banks, Topic tagging, Difficulty levels	Authoring tools, Bulk uploads, Various question types	Detailed analytics, Performance breakdowns, Topic analysis	Team coordination, Grading collaboration	None	AI-powered proctoring, Behavioural detection
Moodle	Categories/subcategories, Versioning, Import/export	Quiz activity integration, Category-based selection	Basic usage statistics, Version history	Course-level sharing, Category management	None	Limited AI features

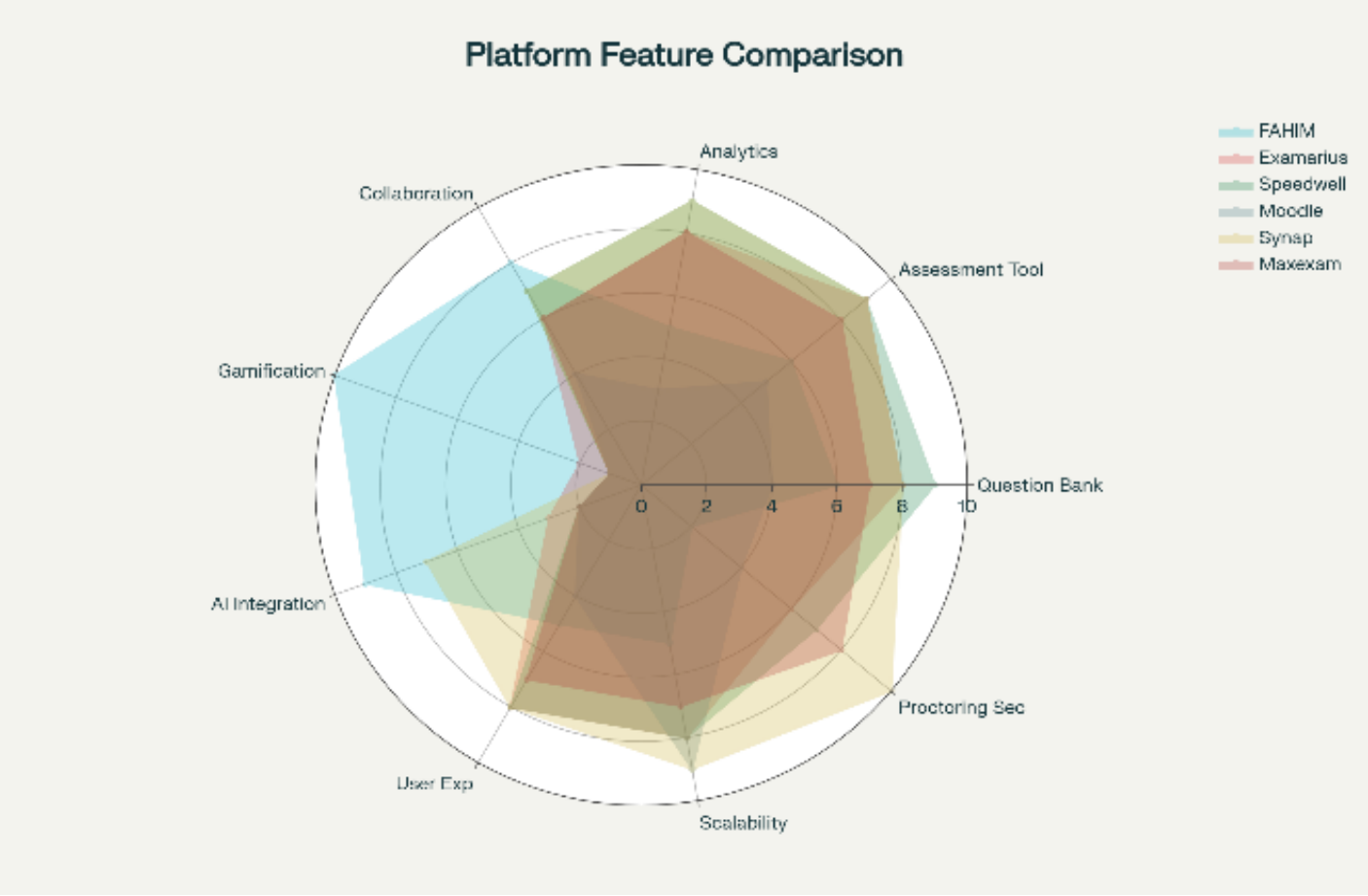
Realistic assessment shows FAHIM trailing in 6 dimensions, particularly as it does not provide proctoring security. While Speedwell, Synap and Examarius provide a complete suite for examination solutions with polished user experience and scalability, FAHIM offers competitive advantages for its unique gamification elements, AI integration and collaboration features in a single platform. FAHIM main purpose is to be a platform for question bank repository rather than a complete examination solution that includes exam proctoring and online test system.

Hence, FAHIM's strength lies not in being the best overall platform, but in being the only platform that successfully gamifies assessment creation while providing AI-assisted pedagogical guidance. This creates a unique niche of question bank application for institutions prioritising:

- Faculty engagement and motivation
- Collaborative assessment development
- Modern, AI-enhanced workflows

Figure 12 illustrates how FAHIM compares across nine critical dimensions against Examarius, Speedwell, Moodle, Synap, and Maxexam.

Figure 12: FAHIM vs. Established Platforms



None of the other platforms offer significant gamification, representing FAHIM's most unique differentiator. While Synap offer AI for proctoring/security, FAHIM is unique in providing AI assistance for content development. This analysis demonstrates that FAHIM offers genuinely unique capabilities that address fundamental challenges in educational assessment not adequately provided by existing platforms.

Potential Findings and Commercialisation

The potential benefits of FAHIM can be divided into three categories: educational enhancement, operational efficiency, and commercial viability.

Educational enhancement

Designing assessments with FAHIM's intuitive interfaces reduces extraneous cognitive load by minimising the need for users to expend mental effort on understanding complex navigation (Brüggemann al., 2022). By employing certain algorithms, it automates parts of the assessment creation process, potentially reducing time spent on mundane tasks such as formatting and alignment checks. Automating routine aspects of assessment creation, such as item generation and analysis, significantly reduces preparation time, freeing educators to focus on refining content quality and alignment (Meylani, 2024).

Tagging questions with Bloom's levels (e.g., remembering, understanding, applying) can help educators ensure a balanced distribution of cognitive complexity across an assessment, improving transparency in what skills are being tested (Kumar, et al., 2024). Furthermore, using CEFR levels to tag language assessment items provides clarity on the linguistic complexity expected, aligning the assessment with standardised language proficiency levels (Gaillat et al., 2021). Tagging learning outcomes and their alignment with curriculum standards can also ensure more consistent alignment with educational goals, providing balanced assessment items (Meylani, 2024). Therefore, collecting and analysing metadata allows educators to gain insights into assessment patterns, student performance, and item effectiveness, thereby enhancing consistency and refinement over time (Kumar et al., 2024; Mathew, 2020). By standardising metadata, the app can also facilitate quicker audits and adherence to quality assurance protocols.

Operational efficiency

Gamification involves integrating game design elements like points, badges, and leaderboards, which can enhance motivation and engagement by appealing to intrinsic and extrinsic motivational factors (Jun & Lucas, 2024; Zeng et al., 2024). These elements available in FAHIM create a sense of achievement and competition, encouraging educators to contribute more actively to shared resources. For example, badges can acknowledge contributions, providing recognition that motivates educators to remain engaged (Imran, 2019). Gamification can enhance the satisfaction and engagement levels of individuals engaging with gamified systems, which in turn can increase participation in community activities such as knowledge sharing (Wang & Wang, 2024). When educators are part of a gamified community, they experience a more dynamic and interactive environment that can motivate them to be more involved.

Leaderboards and public recognition of contributions can drive educators to contribute more to match or exceed their peers' efforts (Jun & Lucas, 2024). Long-term, gamification can lead to a cultural shift where continuous engagement and improvement become embedded in the community's values. Regular interaction and progression within a gamified system foster a sustained contribution, making educators more likely to engage in sharing and developing shared resources over the long term (Wulan et al., 2024). This can lead to an increase in the quantity and quality of shared educational resources, leading to an increase in operational efficiency.

Commercial viability

Market awareness, lack of funding and institutional support are some of the main challenges in commercialising an innovative product (Jamil et al., 2024; Ismail & Sidek, 2019). On paper, this application has potential to be commercialised for several reasons. Firstly, the ideation of this app was born out of necessity where a similar application is not available (at least some of its features). The practical issues raised in this article regarding the challenges in item creation and assessment preparation faced by educators in higher education institutions are universal. FAHIM provides features that are current by considering both international and national frameworks, while integrating with features such as calendar, collaborative tool, gamification element and AI chatbot, further differentiate it from other similar applications. The application of IP (copyright) for this app is also ongoing, which will further enhance its marketability.

Overall, FAHIM has the potential to facilitate improved assessment practices leading to greater efficiency, consistency, and reliability.

Implementation Challenges and Solutions

There are two possible challenges if FAHIM were to be implemented institution-wide, namely technological infrastructure and faculty development.

Firstly, the institution should conduct comprehensive technology audits to identify infrastructure weaknesses and develop upgrade roadmaps aligned with implementation timelines. For hardware compatibility and standards, the use of cloud-based architecture will reduce hardware requirements while ensuring cross-platform compatibility. Using localised servers for FAHIM is recommended but eventually once it has been institutionalised and fully adopted, the server needs to be stress-tested for heavy-user traffic. Hence, successful FAHIM deployment requires addressing connectivity challenges through phased implementation approaches that prioritise reliable, low-bandwidth solution when internet access is inconsistent. Most importantly, implementation must address data security protocols essentials for educational technology adoption. FAHIM's HTTPS encryption and login authentication provide foundational security but could significantly offer more robust security with additional session management and multi-factor authenticator.

From the aspect of faculty development, it requires moving beyond technical training to focus on pedagogical integration of assessment technologies. Training program should emphasise how FAHIM can enhance teaching practices rather than replacing established methods, with particular attention to constructive alignment principles and outcome-based assessment design. It also has to be easily revamped and updated if required, in accordance with the latest development in assessment where there could be a possible evolution from outcome-based to value-based assessment. This requires establishing professional learning communities and ongoing support system, where institutions should implement mentorship programs pairing technology-proficient educators with colleagues, creating a collaborative environment for knowledge sharing and best practice development. Adoption for such technology may take time and faculty resistance to new technology necessitates careful change management strategies addressing concerns about increased workload, learning curves, and perceived threats to teaching autonomy. Voluntary adoption phases, peer champions, and administrative supports for experimentation are crucial and could be the success factors for FAHIM adoption.

NOVELTY AND RECOMMENDATIONS

FAHIM emerges as a pioneering platform designed to transform the process of creating educational assessments by streamlining the intricate task of exam design, ensuring assessments are not only fair but also closely aligned with learning outcomes. Moreover, FAHIM's incorporation of AI chatbot and gamification elements addresses the challenges of engagement and steep learning curves, making assessment design a more interactive and rewarding experience. In this way, FAHIM positions itself as a crucial tool for educators aiming to foster an inclusive and rigorous academic environment, symbolizing a shift from traditional, often hurried exam preparations towards a data-informed and repeatable practice. For it to be empirically evaluated, future research may include a pilot metric indicator such as user satisfaction, assessment quality improvements, and time-saving metrics comparing FAHIM to traditional methods to validate its effectiveness in real-world educational settings.

ACKNOWLEDGEMENTS

I gratefully acknowledge my colleagues whose shared experience and persistent challenges over the years provided the essential inspiration and practical insights that informed the development of this app.

REFERENCES

1. Akintayo, O. T., Eden, C. A., Onyebuchi, N. C., & Ayeni, O. O. (2024). Evaluating the impact of educational technology on learning outcomes in the higher education sector: A systematic review. International Journal of Management & Entrepreneurship Research, 6(5), 1091. https://doi.org/10.51594/ijmer.v6i5.1091
2. Ali, L. (2018). The design of curriculum, assessment and evaluation in higher education with constructive alignment. Journal of Education and E-Learning Research, 5(1), 72–78. https://doi.org/10.20448/journal.509.2018.51.72.78

3. Amir, A. F., Mohamed, A. T. F. S., & Juhari, J. (2024). Sustainable online education for higher education institutions: A systematic literature review. Asia Pacific Journal of Educators and Education, 39(1),45–91. https://doi.org/10.21315/apjee2024.39.1.3
4. Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2000). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.
5. Baharum, N. N., Razali, A. B., Ismail, L., & Nordin, N. (2021). Aligning a university English language proficiency measurement tool with the CEFR: A case in Malaysia. Pertanika Journal of Social Sciences and Humanities, 29(S3), 139–156. https://doi.org/10.47836/pjssh.29.s3.09
6. Balci, S., Secaur, J. M., & Morris, B. J. (2022). Comparing the effectiveness of badges and leaderboards on academic performance and motivation of students in fully versus partially gamified online physics classes. Education and Information Technologies, 27(6), 8669–8704. https://doi.org/10.1007/s10639-022-10983-z
7. Biggs, J., & Tang, C. (2014). Constructive alignment: An outcomes-based approach to teaching anatomy. In New directions in teaching and learning anatomy (pp. 31–44). Springer International Publishing. https://doi.org/10.1007/978-3-319-08930-0_4
8. Bissell, A. N., & Lemons, P. P. (2006). A new method for assessing critical thinking in the classroom. BioScience, 56(1), 66–72. https://doi.org/10.1641/0006-3568(2006)056[0066:anmfac]2.0.co;2
9. Brown, G. T. L. (2022). The past, present and future of educational assessment: A transdisciplinary perspective. Frontiers in Education, 7, Article 1060633. https://doi.org/10.3389/feduc.2022.1060633
10. Brüggemann, T., Ludewig, U., Lorenz, R., & Mcelvany, N. (2022). Effects of mode and medium in reading comprehension tests on cognitive load. Computers & Education, 192, Article 104649. https://doi.org/10.1016/j.compedu.2022.104649
11. Cheng, L., Hu, H., & Rogers, T. (2004). ESL/EFL instructors' classroom assessment practices: Purposes, methods, and procedures. Language Testing, 21(3), 360–389. https://doi.org/10.1191/0265532204lt288oa
12. Deterding, S. (2015). The lens of intrinsic skill atoms: A method for gameful design. Human–Computer Interaction, 30(3–4), 294–335. https://doi.org/10.1080/07370024.2014.993471
13. Gaillat, T., Sousa, A., Zarrouk, M., Ballier, N., Simpkin, A., Bouyé, M., & Stearns, B. (2021). Predicting CEFR levels in learners of English: The use of microsystem criterial features in a machine learning approach. ReCALL, 33(3), 1–18. https://doi.org/10.1017/s095834402100029x
14. Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. European Journal of Education, 57(4), 542–570. https://doi.org/10.1111/ejed.12533
15. Imran, H. (2019). Evaluation of awarding badges on student's engagement in gamified e-learning systems. Smart Learning Environments, 6, Article 9. https://doi.org/10.1186/s40561-019-0093-2
16. Ismail, N., & Sidek, S. (2019). Determinant factors for commercialising research products in Malaysian public universities. International Journal of Innovative Technology and Exploring Engineering, 8(6S4), 780–787. https://doi.org/10.35940/ijitee.f1157.0486s419
17. Jamil, A. A., Indiran, L., Roslan, S., Lim, S. A. R., Ismail, K., & Muhamad, S. (2024). Surmounting commercialization barriers in Malaysian research universities: Paving the way for unprecedented innovation impact. International Journal of Research and Innovation in Social Science, 8(11), 682–698. https://doi.org/10.47772/ijriss.2024.8110054
18. Kang, H., & Furtak, E. M. (2021). Learning theory, classroom assessment, and equity. Educational Measurement: Issues and Practice, 40(3), 73–82. https://doi.org/10.1111/emip.12423
19. Kerneža, M., & Zemljak, D. (2023). Science teachers' approach to contemporary assessment with a reading literacy emphasis. Journal of Baltic Science Education, 22(5), 851–864. https://doi.org/10.33225/jbse/23.22.851
20. Kumar, V., Chandrappa, & Harinarayana, NS. (2024). Exploring dimensions of metadata quality assessment: A scoping review. Journal of Librarianship and Information Science. https://doi.org/10.1177/09610006241239080
21. Lee, N. A. A., Kassim, A. A. M., & Bakar, R. A. (2022). The CEFR-aligned curriculum execution in Malaysia and other countries: A conceptual paper. Malaysian Journal of ELT Research, 19(1), 1–15. https://doi.org/10.52696/tgct6849
22. Luo, J. (2024). Validating the impact of gamified technology-enhanced learning environments on motivation and academic performance: Enhancing TELEs with digital badges. Frontiers in Education, 9, Article 1429452. https://doi.org/10.3389/feduc.2024.1429452

23. Mathew, A. C. (2020). Prediction of question tags based on LDA and deep neural network. International Journal of Computing, Communications and Networking, 9(2), 7–10. https://doi.org/10.30534/ijccn/2020/02922019
24. Meylani, R. (2024). A comparative analysis of traditional and modern approaches to assessment and evaluation in education. Batı Anadolu Eğitim Bilimleri Dergisi, 15(2), 1–17. https://doi.org/10.51460/baebd.1386737
25. O'Neill, G. (2017). It's not fair! Students and staff views on the equity of the procedures and outcomes of students' choice of assessment methods. Irish Educational Studies, 36(2), 221–236. https://doi.org/10.1080/03323315.2017.1324805
26. Perry, N. E., Calder, K., & Walton, C. (1999). Teachers developing assessments of early literacy: A community of practice project. Teacher Education and Special Education: The Journal of the Teacher Education Division of the Council for Exceptional Children, 22(4), 255–270. https://doi.org/10.1177/088840649902200404
27. Prajapati, M. (2022). Introductory guide to the Common European Framework of Reference (CEFR) for English language teachers. Integrated Journal for Research in Arts and Humanities, 2(6), 291–296. https://doi.org/10.55544/ijrah.2.6.40
28. Ragupathi, K., & Lee, A. (2020). Beyond fairness and consistency in grading: The role of rubrics in higher education. In Diversity and inclusion in global higher education (pp. 33–54). Springer Singapore. https://doi.org/10.1007/978-981-15-1628-3_3
29. Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. WIREs Data Mining and Knowledge Discovery, 10(3), Article e1355. https://doi.org/10.1002/widm.1355
30. Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. American Psychologist, 55(1), 68–78. https://doi.org/10.1037/0003-066x.55.1.68
31. Sclater, N., & MacDonald, M. (2004). Putting interoperability to the test: Building a large reusable assessment item bank. Research in Learning Technology, 12(3), 225–237. https://doi.org/10.3402/rlt.v12i3.11254
32. Smith, A., Mcconnell, L., Iyer, P., Allman-Farinelli, M., & Chen, J. (2024). Co-designing assessment tasks with students in tertiary education: A scoping review of the literature. Assessment & Evaluation in Higher Education, 49(6), 1–17. https://doi.org/10.1080/02602938.2024.2376648
33. Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. Learning and Instruction, 4(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5
34. Szwarc, E., Bach-Dąbrowska, I., & Bocewicz, G. (2018). Competence management in teacher assignment planning. In Advances in intelligent systems and computing (Vol. 835, pp. 449–460). Springer International Publishing. https://doi.org/10.1007/978-3-319-99972-2_37
35. Uri, N. F. M. (2023). Challenges in CEFR adoption: Teachers' understanding and classroom practice. International Journal of Modern Languages and Applied Linguistics, 7(1). https://doi.org/10.24191/ijmal.v7i1.7522
36. Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. Decision Sciences, 39(2), 273–315. https://doi.org/10.1111/j.1540-5915.2008.00192.x
37. Viberg, O., Mutimukwe, C., Hrastinski, S., Cerratto-Pargman, T., & Lilliesköld, J. (2024). Exploring teachers' (future) digital assessment practices in higher education: Instrument and model development. British Journal of Educational Technology, 55(6), 2597–2616. https://doi.org/10.1111/bjet.13462
38. Volkov, S. A. (2023). Applying the cognitive congruence principle to target language training. Training, Language and Culture, 7(4), 54–67. https://doi.org/10.22363/2521-442x-2023-7-4-54-67
39. Wang, X.-W., & Wang, Z. (2024). The influence of gamification affordance on knowledge sharing behaviour: An empirical study based on social Q&A community. Behaviour & Information Technology, 43(10), 1–17. https://doi.org/10.1080/0144929x.2024.2361059
40. Wisniewski, K. (2017). Empirical learner language and the levels of the Common European Framework of Reference. Language Learning, 67(S1), 232–253. https://doi.org/10.1111/lang.12223
41. Wulan, D. R., Fiyul, A. Y., Hidayat, Y., Nainggolan, D. M., & Rohman, T. (2024). Exploring the benefits and challenges of gamification in enhancing student learning outcomes. Global International Journal of Innovative Research, 2(7), 238. https://doi.org/10.59613/global.v2i7.238

42. Zaidi, N. L. B., Monrad, S. M., Kurtz, J. B., Tai, A., Santen, S. A., Gruppen, L. D., Ahmed, A. Z., & Grob, K. L. (2018). Pushing critical thinking skills with multiple-choice questions: Does Bloom's taxonomy work? Academic Medicine, 93(6), 856–861. https://doi.org/10.1097/acm.0000000000002087
43. Zeng, J., Fan, A. C. W., Looi, C.-K., & Sun, D. (2024). Exploring the impact of gamification on students' academic performance: A comprehensive meta-analysis of studies from the year 2008 to 2023. British Journal of Educational Technology, 55(6), 1–22. https://doi.org/10.1111/bjet.13471