SCIENTIFIC Publishing Limited

Digital Technologies Research and Applications

http://ojs.ukscip.com/index.php/dtra

Article

Designing and Governing Trustworthy AI Marketing Systems in Educational Technology: A Managerial and Implementation Framework

Chuntie Chen [1,*], Nor Hidayati Binti Zakaria [1], Wei Deng [2], Xiaoli Xu [3] and Youyu Xu [1]

[1] Azman Hashim International Business School, Universiti Teknologi Malaysia (UTM), Kuala Lumpur 56100, Malaysia

[2] School of Economics and Management, Huizhou University, Huizhou 516000, China

[3] Department of Accounting, Guizhou Electronic Commerce Vocational College, Guiyang 550000, China

[*] Correspondence: chenchuntie@mail.com

Received: 9 January 2026; **Revised:** 11 February 2026; **Accepted:** 19 March 2026; **Published:** 24 April 2026

Abstract: This study proposes a dual-dimensional framework to design such technologies in the EdTech industry using the conceptual framework methodology, which is based on thematic content analysis and theory-driven and evidence-based validation using scholarly literature, industry reports, and policies from the top EdTech organizations worldwide. The framework identifies trustworthiness as the key construct that integrates two interrelated dimensions. The governing dimension is operational at the managerial level and includes governance principles based on an extended Unified Theory of Acceptance and Use of Technology (UTAUT) framework and personalization governance principles based on three dimensions, such as intensity, tempo, and boundaries. The designing dimension is operational at the implementation level and identifies the technical requirements related to virtual sales personnel systems and AI promotion systems. The results of the validation process against known platform practices demonstrate a mixed pattern of alignment, with stronger support in regulatory-related areas and weaker support in governance-intensive domains where the framework extends current industry practice. The extension of the UTAUT model from personal acceptance to organizational governance represents a theoretical contribution with a link to existing research in personal acceptance and its expanded applicability. The three-dimensional personalization governance model has more detailed mechanisms than the current one-dimensional approaches. For educational technology organizations, the framework offers systematic guidance in the development of AI-based marketing systems that are trustworthy to users while being effective for organizational goals. Validation results indicate tempo governance and recommendation explainability as areas that need development in the industry for effective engagement with users in a sustainable manner.

Keywords: AI Marketing Systems; Educational Technology; Trustworthiness; Governance Framework; UTAUT Extension

1. Introduction

The rapid development of artificial intelligence (AI) has significantly impacted marketing trends in various fields, with educational technology being a notable area where AI marketing innovations are exhibiting dynamic trends. An analytical study of AI marketing reveals that intelligent systems have successfully replaced human marketers in analytical, intuitive, and empathetic marketing roles, thereby transforming customer engagement and value creation paradigms [1]. This evolution is not only technological, as studies on AI from a multidisciplinary

perspective reveal the new challenges and opportunities it presents, thus requiring a comprehensive approach to its benefits while addressing the associated risks [2]. The value of the education technology industry exceeds $400 billion, and the pace of adoption of AI-based marketing systems such as virtual sales staff and recommendation systems has accelerated.

In this context of change, the applications of artificial intelligence in marketing have shown significant promise in improving customer relationships, and literature-based analyses of the applications of this field reveal that content personalization, customer segmentation, and conversational agents are the key application domains [3]. The interplay of AI and higher education has been the focus of considerable attention in scholarship, as systematic reviews highlight the wide range of applications, including intelligent tutoring systems, administration, and student services [4]. Contemporary assessments of the use of AI in an educational context highlight the move beyond the experimental to the mainstream, including the use of virtual sales personnel and AI-based promotional systems in how the educational system interfaces with current and prospective students [5]. These developments emphasize the importance of realizing the implications of effectively using AI-driven marketing systems by educational technology organizations for their business and educational goals.

Previous studies have mainly examined existing AI-based marketing systems from user perspectives, with emphasis on the trust formation process and technology acceptance issues. Another area that research has focused on is AI-based personalization in interactive marketing environments, where studies have sought to investigate how this interactive personalization impacts the value, eventually culminating in behavioral intentions [6]. Research on the importance of customer trust in AI systems has shown anthropomorphism, empathic response, and interaction quality as key antecedents, highlighting the significance of AI in marketing interface design in creating customer trust perception [7]. Research carried out using the Technology Acceptance Model (TAM) frameworks to investigate AI adoption in higher educational institutions found performance expectancy, effort expectancy, and social influence to be key predictors of behavioral intention to use AI in educational tools [8]. Research extending the TAM to measure the intention to use AI-based learning platforms by university students found that perceived usefulness, perceived ease of use, and facilitating conditions were significant, which is in line with the extension of acceptance model use in general [9]. However, these studies seem to be focusing more on individual acceptance and trust-building aspects rather than design and governance issues.

While considerable progress has been made in the implementation of AI-based marketing systems in educational technology, there are gaps in the understanding of the ways in which organizations can develop such systems in order to ensure trustworthiness. Earlier research on the use of AI in educational settings has identified the basic concerns related to the choices, autonomy, and the rights of the learners and has emphasized the need to address the ethical concerns related to the use of such systems, which have not been sufficiently explored in the literature [10]. Studies of large language models in educational settings have identified opportunities and challenges related to artificial intelligence. Although there is unprecedented promise in personalization and student engagement with AI, accuracy and proper usage of AI have highlighted the importance of effective governance arrangements in translating ethical principles into practice [11]. From the analysis of the digital transformation in the field of AI, the effective implementation is identified as a major challenge since there are difficulties in the implementation of the opportunities offered by AI through the lack of effective frameworks linking the strategic intent and implementation [12]. These identified gaps demonstrate a significant disparity existing between user-level mechanisms and organizational implementation practices, highlighting the importance of integrative frameworks that consider design specifications and governance principles in trustworthy AI marketing systems in educational technology.

The theory base for bridging the gap extends and adds to the existing Unified Theory of Acceptance and Use of Technology (UTAUT), which has established acceptable constructs for understanding the adoption of technology in terms of performance, effort, social factors, and facilitating conditions [13]. An analysis of the customer experience in virtual environments reveals the involvement of AI systems in influencing customer engagement at various points and hence the need for the organization's governance structure to consider the full range of such interactions as opposed to specific parts of the system [14]. Studies on explainability and its relationship to trust have also shown that the level of trust that users have in AI systems is dependent on the level of understanding that the user has regarding the system's functioning. This, in turn, emphasizes the need for ensuring that transparency is integrated in the system's governance principles [15]. This also includes the technical specifications, considering the relationship between trust and the system's decision rationales' accessibility [16]. Based on the theoretical

underpinnings discussed above, this study extends the UTAUT model from individual user acceptance to organizational governance, where performance expectancy is related to value alignment principles, effort expectancy to user autonomy principles, social influence to transparency principles, and facilitating conditions to design standards.

This model addresses the call for integrative governance of AI-enabled education [17], employing the two-dimensional framework introduced in this study. Furthermore, the framework meets the need for the translation of the impact of AI on education practice and policy into organizational practice through the provision of a systematic approach that balances strategic governance and technical implementation [18]. The proposed framework designates trustworthiness as the unifying attribute of governance principles and design specifications, which is in accordance with international policy recommendations that focus on human-centric AI development in education [19]. A significant innovation in the proposed framework is the three-dimensionality of the governance approach to personalization, which involves intensity, tempo, and boundary considerations. This extends beyond existing frameworks for applying chatbots, which discuss the potential of AI systems but fail to reveal the underlying structures that are in use [20]. The design specifications for virtual sales staff are developed from existing research on AI student assistants, which translates existing knowledge on effectively applying AI in education to design requirements [21]. The implications of this study are significant in that it extends existing literature on the impact of virtual teaching assistants by providing guidelines for organizations on how to effectively use AI-driven marketing systems without compromising educational goals [22]. The two-dimensional approach to managerial governance and system design can provide an educational technology organization with a complete blueprint to design an AI marketing system that not only motivates user trust but also meets the organizational objectives.

2. Methodology

2.1. Research Design

The research utilizes a conceptual framework development approach in developing a dual-dimensional conceptual framework for the design and governance of trustworthy AI marketing systems in educational technology platforms. Conceptual framework analysis is based on a systematic methodology for synthesizing interrelated concepts into coherent theoretical structures, prioritizing the networked nature of concepts rather than linear causality [23]. The methodology is well-suited to the research objectives, as the convergence of AI marketing systems, educational technology, and organizational governance is a new field of study. Inconsistency in the findings of user-level trust, technology acceptance, and AI governance necessitates a conceptual integration of the knowledge gap between the validation of the mechanisms and the implementation of the organizations. The conceptual framework is appropriate as it can capture the complexity in the development of trustworthy AI systems, which involves different stakeholders, different components, and different levels of complexity in the organizations. Such conceptual frameworks are therefore methodologically adequate and theoretically productive in research domains where the necessary constructs are still underdeveloped. This study directly addresses this issue, since the necessary governance constructs at the organizational level for trustworthy AI marketing systems are still undeveloped, despite the availability of user-level empirical research on trust and technology acceptance. In this way, the dual-dimensional framework provides a theoretical architecture in which future research studies, carried out at the level of the organization, surveys, or experimental deployments, can be conducted.

The research design recognizes trustworthiness as the essential organizing attribute with respect to which the principles of governance and specifications of design cohere, reflecting the dual emphasis on "designing" and "governing" evident in the research title. Trustworthiness in AI marketing systems is not built by individual technical constructs or policy statements, but by their integration at the intersection of organizational governance and implementation. The theoretical framework extends UTAUT by adding organizational acceptance, which is achieved by conducting a systematic mapping of the relevant constructs. Specifically, the value alignment principles have been included in the performance expectancy construct to ensure that the AI marketing is aligned with the educational goals. The effort expectancy construct has included the user autonomy principles, and the social influence construct has also included the transparency principles. The facilitating conditions cover the development of implementation standards, which specify the technical requirements for reliable implementation. The framework operationalizes the extension of the theory through the introduction of governance mechanisms at the managerial level and implementation standards at the technical level, thus providing direction for organizations in the education technology

field.

2.2. Data Sources

The framework development incorporated information from three different categories of information sources in terms of their theoretical, practical, and policy aspects. A literature search was conducted through Scopus (https://www.scopus.com/) and Web of Science (https://www.webofscience.com/) using the search string "(artificial intelligence OR AI OR machine learning) AND (educational technology OR EdTech OR e-learning) AND (marketing OR personalization OR recommendation) AND (trust OR governance OR ethics)" covering publications from 2019 to 2024. From the overall dataset of 1,247, a total of 86 peer-reviewed articles met the inclusion criteria. These articles focused on the application of AI in educational settings, which includes marketing, personalization, and governance-related applications. They also provided the necessary theoretical constructs for the conceptual architecture of the framework. The data collection process was carried out between January and March 2024. In total, there were 86 articles that helped in the development of the conceptual framework, with 33 of these articles directly referenced in the manuscript to support the theoretical arguments and components of the conceptual framework as a whole.

The reports used in this study were obtained from credible market intelligence companies. Data on the EdTech market and trends in AI adoption were obtained from HolonIQ Global EdTech Intelligence (https://www.holoniq.com/), including the 2024 Global EdTech Market Report projecting industry growth trajectories. McKinsey Global Institute reports (https://www.mckinsey.com/mgi/) and Deloitte AI Institute publications (https://www.deloitte.com/ai-institute/) provided additional information on the challenges of AI implementation and the trends of organizational transformation for the period 2020–2024. These industry publications provided additional information on the background of the market dynamics, emerging business models, and implementation barriers.

Policies from the platform of three premier educational technology organizations were used as documentary evidence for the validation of the proposed framework. The privacy policies, terms of service, and other documentation were collected from Duolingo (https://www.duolingo.com/privacy), Coursera (https://www.coursera.org/about/privacy), and Khan Academy (https://www.khanacademy.org/about/privacy-policy), with current versions accessed in March 2024. The platforms were chosen based on three criteria. First, market leadership in distinct markets: language learning, professional education, and K-12 instruction, respectively. Second, the implementation of AI marketing systems, such as virtual sales staff and recommendation engines. Third, publicly accessible documentation of governance models. Triangulation of the sources ensured the validity of the framework through alignment with existing theory, contemporary market conditions, and industrial implementation. The information provided is from publicly available documents and does not require ethical approval from the institution.

2.3. Analysis Process

This study adopted a three-phase analytical approach, which combined thematic analysis, theory-driven framework development, and validation through evidence-based methods [24,25]. In Phase I, concept identification was conducted through progressive coding strategies, which were applied to all data sets. Open coding yielded initial concepts related to AI marketing governance and trustworthy system design, such as transparency, user control, personalization, and technology. Axial coding was then used to establish a relationship between governance and design factors. Finally, selective coding integrated the concepts into understandable categories that matched the two-dimensional construct.

The second phase involved framework integration, in which the concepts were grouped based on the expanded UTAUT framework. The principles of governance—transparency, user autonomy, and value alignment—were linked to the managerial level, whereas the specifications related to the virtual sales personnel and the AI promotion systems were linked to the implementation level. The trustworthiness factor acted as the bridging element in the two dimensions. In this phase, a personalization governance model was created, which includes the intensity, tempo, and boundary dimensions of personalization. In addition, the relationship between abstract governance principles and design specifications is instantiated in detail.

The third phase was validation, where the elements of the framework were mapped onto the existing practices of the Duolingo, Coursera, and Khan Academy platforms. In this phase, the level of alignment between the proposed elements and the existing implementations was evaluated. The alignment was evaluated based on the governing and designing dimensions, using a three-tier classification system that included full implementation, partial imple-

mentation, and not applicable. Overall, the three phases represented an iterative process, in which the validation results were used to modify the results of the preceding phases.

To further ensure the analytical rigor, a secondary reviewer was utilized to analyze a representative sample of the data corpus using the same process of open, axial, and selective coding. Any discrepancies in the coding process were resolved through deliberation and re-examination of the original source materials. In addition, triangulation of the three types of source materials was performed to ensure that each of the dominant construct components was supported by evidence from at least two of the source types, thereby reducing the likelihood that the constructs reflect the peculiarities of the data type rather than the underlying properties of the phenomenon of interest. The research process is illustrated in **Figure 1**.

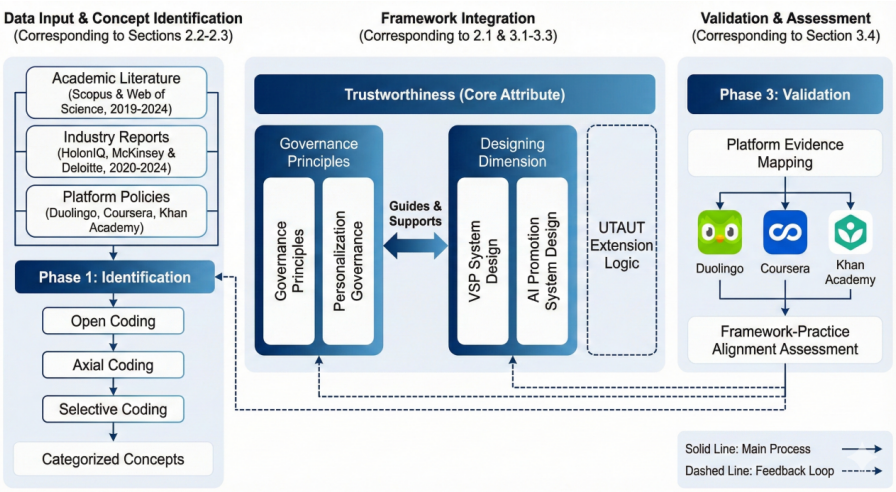

Figure 1. Research Design and Analysis Process.

As depicted in **Figure 1**, the three-phase approach is used. Concepts are progressively coded and categorized based on the findings of the scholarly literature, industry reports, and platform policies. Concepts are organized in a dual-dimensional framework with trustworthiness as the primary attribute, and validation is conducted using the concepts of evidence mapping on the platform.

3. Results

3.1. Resulting Framework Architecture

The development of the conceptual framework yielded a two-dimensional structure that brings together managerial governance and implementation design in support of trustworthy AI marketing systems in educational technology platforms. The architecture of the framework is consistent with the research goal of integrating user-level trust mechanisms and organizational-level implementation practices. The attribute of trustworthiness serves as the unifying concept that brings together governance principles and design specifications in the two dimensions. Achieving trustworthy AI marketing in the context of educational technology platforms cannot be accomplished by either policy initiatives or technical capabilities in isolation. Rather, trustworthy marketing requires the integration of the two, in which the capabilities of governance constrain and inform the capabilities of implementation.

The framework identifies two interrelated dimensions, which are concerned with different organizational levels and functional issues. The first is the governing dimension, which is concerned with the managerial level. It includes the principles of governance based on the extension logic of UTAUT and the principles of personalization governance, which transform ethical principles into management practices. The second is the design dimension, which operates at the implementation level. It specifies the technological needs of virtual sales personnel systems and AI-based promotion systems, which transform the principles of governance into management practices. The relationship between the two dimensions is mutual.

Trustworthiness also functions differently across the two dimensions. While it serves as the goal and the con-

straint at the managerial level, at the implementation level, trustworthiness is the standard for the design. At the managerial level, trustworthiness represents the organizational commitment that the governance principles seek to achieve, as well as the boundary condition for marketing practices. At the implementation level, trustworthiness is specified as a set of design requirements for system developers to ensure alignment with organizational governance principles. This is the gap in the existing literature that the proposed framework addresses. Previous studies have established the importance of trust in the marketing effectiveness of AI-driven systems, but the question of how this trust is achieved has remained unanswered. **Figure 2** illustrates the entire framework architecture.

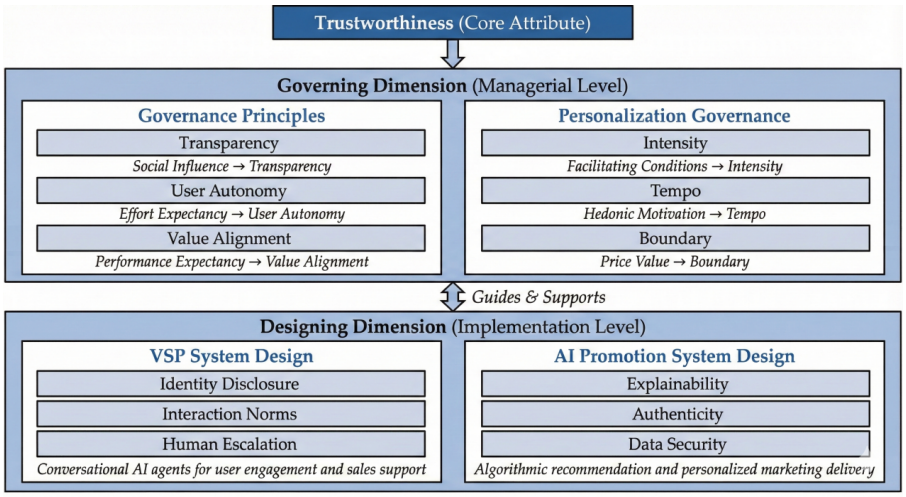

Figure 2. Dual-Dimensional Framework for Trustworthy AI Marketing Systems in EdTech.

Note: Legend ↕ Bidirectional relationship: governance guides implementation; implementation informs governance feasibility.

As illustrated in **Figure 2**, the dual-dimensional model features trustworthiness as the key attribute. The governing dimension comprises governance principles derived from the fundamental constructs of UTAUT: performance expectancy, effort expectancy, and social influence, as well as personalization governance mechanisms, which extend UTAUT2's consumer-oriented concepts of facilitating conditions, hedonic motivation, and price value to organizational governance. The designing dimension specifies the design of VSP systems for conversational AI agents and AI promotion systems for algorithmic recommendations. These two dimensions have bidirectional relationships with each other, where governance principles guide technical implementation decisions and capabilities, and vice versa.

3.2. Governing Dimension: Managerial Level Components

The governing dimension is composed of two different components. Governance principles are the basic commitments that inform all AI marketing activities. Personalization governance mechanisms are the fine-tuned control measures that are used to manage AI-based personalization in the educational setting.

The transparency principle is associated with the degree of visibility of the algorithmic decision-making process and the prerogative of users. This principle is underpinned by the social influence construct of the UTAUT model. Users' perception of system trustworthiness depends upon their ability to understand the system's operations and the underlying logic of the recommendation. Operationally, the transparency principle requires marketers to reveal the use of AI in marketing communications, explain the logic of the recommendation in simple terms, and provide users with insights into the data factors that influence the content of the marketing communications. The principle of user autonomy relates to the level of control that the learner has in relation to the utilization of personalization parameters and data utilization. This principle is a result of the effort expectancy construct. Ease of use in AI interaction is paramount in ensuring the acceptance and trust of the users. Therefore, it is crucial for the organization to ensure that it provides enough configurations, opt-in or opt-out features, and the ability to limit AI-based marketing interactions while ensuring access to the educational services. The value alignment principle deals with the alignment of AI marketing objectives with the educational mission. This principle is based on the

performance expectancy construct, which states that the value generated by the AI system must be aligned with the educational objectives of the learner rather than conversion objectives.

The personalization governance component proposes the triadic personalization governance framework, which replaces the unidimensional personalization intensity construct. The personalization intensity dimension deals with the extent and rate of marketing communication generated by AI and is based on the facilitating conditions construct, which states that organizational and technological infrastructures determine the extent and rate of communication that can be potentially manageable and monitorable. In terms of operations, organizational needs include setting limits on the number of messages, monitoring fatigue, and implementing progressive exposure based on the response to the messages. The tempo dimension relates to the coordination between the timing of marketing activities and the learning process. The theory that guides this dimension is the hedonic motivation construct, which states that the user experiences satisfaction and happiness through the hedonic value that the educational website offers due to uninterrupted learning. The interruption, which is the promotional communication, affects the hedonic value of the marketing campaign and hence the satisfaction level of the user. From the operational perspective, the distribution of promotional materials would be in line with the educational cycles so as not to interfere during specific periods when the users are in concentrated learning sessions. It would be during the natural phases when the users are more open to the promotional messages. The boundary dimension specifies the scope of the use and interaction of the data in relation to the AI marketing. This dimension utilizes the price-value framework, in which users constantly compare the value of personalized services in terms of the costs of providing their data to marketers. When marketers exceed the boundaries of data usage for marketing, it affects the value of the platform to the user. The operational consideration includes classifying the data based on the sensitivity level, securing the information categories deemed inappropriate for marketing purposes, and specifying the functional boundaries between the two functions. The specifications for the managerial level component are shown in **Table 1**.

Table 1. Managerial Level: Governance Principles and Personalization Governance.

Component	Definition	UTAUT Origin	Core Mechanisms	Platform Evidence
Governance Principles				
Transparency	Visibility of algorithmic decision-making processes and users' right to understand AI system operations	Social Influence	AI involvement disclosure; Recommendation logic explanation; Data factor visibility	Coursera: explicit recommendation rationale; Duolingo: general algorithmic disclosure; Khan Academy: data usage transparency
User Autonomy	Learners' control over personalization parameters and data usage in AI marketing interactions	Effort Expectancy	Preference settings provision; Opt-in/opt-out mechanisms; Modification without service sacrifice	Duolingo: granular notification preferences (Full); Coursera: broader but less detailed control options (Partial); Khan Academy: comprehensive account and privacy controls (Full)
Value Alignment	Consistency between AI marketing objectives and educational mission priorities	Performance Expectancy	Educational goal prioritization; Commercial restraint mechanisms; Learning outcome protection	Khan Academy: strong mission alignment (non-profit); Duolingo: partial balance; Coursera: ongoing optimization tension
Personalization Governance				
Intensity	Frequency and density management of AI-generated marketing communications	Facilitating Conditions	Upper limits on promotional volume; Fatigue monitoring implementation; Progressive exposure strategies	Duolingo: push notification limits; Coursera: email frequency caps; Khan Academy: minimal promotional intensity
Tempo	Timing alignment between marketing communications and learning cycles	Hedonic Motivation	Learning rhythm respect; Disruption avoidance during focus periods; Transition moment synchronization	Limited implementation across platforms; Duolingo: learning reminder timing only (Partial); Coursera: no documented tempo mechanisms (Partial); Khan Academy: no systematic implementation (Partial); Area for industry development
Boundary	Scope definition of data usage and interaction contexts for AI marketing	Price Value	Data sensitivity classification; Information category protection; Functional boundary establishment	All platforms: policy-level commitment; Coursera: detailed data categorization; Operational mechanisms vary

Notes: The UTAUT extensions for personalization governance dimensions represent conceptual mappings rather than direct construct transfers. Intensity extends from facilitating conditions through the logic that organizational infrastructure determines manageable communication volumes. Tempo extends from hedonic motivation through the recognition that timing affects user enjoyment of platform engagement. The boundary extends from the price value through the user's cost-benefit evaluation of data disclosure for personalization benefits.

Table 1 summarizes the governing dimension components. It presents three governance principles with corre-

sponding UTAUT origins and the three-dimensional personalization governance model. Platform evidence reveals variable implementation patterns. Tempo governance shows limited adoption across all examined platforms, indicating an area for continued development.

3.3. Designing Dimension: Implementation Level Specifications

The design dimension outlines the technical specifications for two different types of AI marketing systems. First, virtual sales personnel systems will interact with consumers using conversational AI technologies, whereas AI promotion systems will offer personalized recommendations using algorithmic content selection. Second, these specifications will translate the design principles of AI marketing systems into technical specifications that development teams can actually implement. Thus, the technical architecture and systems must align with the trustworthiness promises identified in the governing dimension.

The design of virtual sales personnel systems has three specification domains, all of which pertain to trust-related concerns specific to AI agents that simulate human-like sales dialogues in educational settings. The specification domains are: the identity disclosure specification, which calls for the explicit statement of the artificial nature of agents in all interactions within the VSP system. This ensures that users are aware that they are indeed interacting with artificial agents rather than human agents, thus preventing any deceptive practices that could undermine trust. The interaction norms specification defines the boundaries of the content that is to be shared in VSP communications. In an educational setting, improper marketing pressure has the potential to adversely affect the well-being and educational outcomes. In this regard, the interaction norms specification emphasizes the necessity of content moderation systems that are effective in filtering out improper content. The human escalation specification defines the conditions for escalating VSP communications to human agents. However, there are circumstances that require human judgment, emotion, and authority that extend beyond the capabilities of the AI system. Such cases as complex investigations that go beyond the capabilities of the VSP system, emotional distress, and complaints represent the circumstances that need to be addressed by the human escalation specification.

The AI promotion system design specification has three domains. These domains ensure that the algorithmic recommendation and targeting system works in accordance with the set criteria for trustworthiness. The explainability specification requires that the rationale for the recommendation be easily understood by the users, thus allowing the learners to understand the basis for the recommendation of the course or feature and the factors that influence the content personalization. Two implementation approaches are offered as examples of how this specification might be achieved: a factor-disclosure interface where the user is shown the factors that are most relevant to the current recommendation—such as previous course history and learning objectives—in terms of how they contribute to the recommendation; and a counterfactual adjustment approach where users are able to modify their preference parameters and see in real-time how the recommendations are affected, thus making the logic of the recommendation engine transparent through interactive means rather than simply disclosing it. The authenticity specification is related to the accuracy of the information provided in the AI-generated promotional material. It requires the provision of mechanisms for verification of the accuracy of the information provided in the promotional material with regard to the accuracy of the courses offered. The data security specification defines the requirements for protecting user data in the AI marketing system, which includes the encryption of data and the provision of an audit trail for monitoring data usage. The comprehensive implementation specification is shown in **Table 2**.

Table 2 presents the specifications in the design dimension based on the VSP system design and the AI promotion system design. Evidentiary data from the platform has shown better alignment in the data security specifications. Further development is required in the areas of explainability and VSP identity disclosure.

3.4. Framework Validation

The proposed framework is examined against the existing practices of three prominent educational technology platforms in three different criteria: full implementation, partial implementation, and non-applicable, as discussed in the following section.

Validation of the governing dimension shows mixed alignment across platforms, with notable discrepancies related to divergent organizational priorities and system maturity. Transparency features are observed in the documentation of all platforms, although the degree of transparency differs significantly across platforms. Features related to autonomy show heterogeneous platform implementation, although Duolingo shows complete alignment

in notification preferences. Account management and privacy were fully provided by Khan Academy, leading to full alignment. Coursera offered a broader range of, yet less granular, control mechanisms, resulting in partial alignment. The highest divergence in value alignment was noted between the platforms. With personalization governance, the extension of the framework beyond existing practice is revealed. Regarding intensity governance, the application of frequency limits is partially evident in commercial platforms. There is a lack of fatigue monitoring, with the minimal promotion strategy of Khan Academy showing full alignment. There is a lack of development in tempo governance in all platforms, relating to learning cycle synchronization. Boundary governance exists at a policy level but without defined mechanisms.

Table 2. Implementation Level: System Design Specifications.

System Type	Design Element	Definition	Specification Requirements	Platform Evidence
VSP System Design	Identity Disclosure	Clear identification of AI nature in all virtual sales personnel interactions	Prominent AI labeling in chat interfaces; Explicit disclosure statements at conversation initiation; Visual design elements distinguishing AI from human channels	Khan Academy (Khanmigo): explicit AI identification with prominent labeling; Coursera: partial disclosure mechanisms; Duolingo: no comparable VSP functionality currently deployed
	Interaction Norms	Content boundaries and behavioral guidelines for VSP communications in educational contexts	Content moderation systems filtering prohibited topics; Response guidelines maintaining professional boundaries; Escalation protocols for sensitive interactions	Khan Academy: comprehensive content filtering and teacher-appropriate guidelines; Coursera: basic moderation systems; Industry-wide variation in norm specificity
	Human Escalation	Trigger conditions and handoff procedures for transferring interactions from AI to human agents	Complex query detection exceeding VSP capabilities; Emotional distress indicator recognition; Formal complaint routing protocols; Seamless transition mechanisms	All platforms: customer support channels available; Khan Academy: teacher escalation protocols; Coursera: help center integration; Response time varies significantly
AI Promotion System Design	Explainability	Comprehensibility of recommendation rationale enabling users to understand personalized content selection	User-facing explanation of recommendation factors; Accessible rationale presentation; Transparency in data influence on suggestions	Coursera: partial recommendation rationale for selected course suggestions; Duolingo: limited explanation for lesson sequencing; Khan Academy: minimal user-facing explainability; Area requiring continued development across all platforms
	Authenticity	Information accuracy in AI-generated promotional content ensuring claims represent actual offerings	Verification mechanisms for course and outcome claims; Source attribution for third-party information; Regular accuracy audits	All platforms: content review processes established; Coursera: detailed course information verification; Quality assurance maturity varies across platforms
	Data Security	Protection requirements for user data employed in AI marketing systems	Encryption standards for transmission and storage; Access control limiting data availability; Audit trail capabilities for usage review	All platforms: robust implementation reflecting regulatory compliance; GDPR and CCPA alignment documented; Industry-leading security practices consistent

Notes: Duolingo's VSP System Design is marked as N/A because, at the time of data collection (March 2024), Duolingo's AI features (such as Duolingo Max's Roleplay and Explain My Answer) functioned primarily as learning support tools rather than marketing-oriented virtual sales personnel systems. These features do not perform sales, subscription promotion, or customer acquisition functions that would qualify them as VSP systems under this framework's definition.

The validation of the design dimension revealed heterogeneous implementation patterns, as expected, due to the diverse maturity levels of AI marketing technologies in modern EdTech practice. The specifications of the VSP system presented the largest implementation gaps, as conversational AI marketing is still an evolving domain. The specifications for the AI promotion system indicate that the area of explainability needs further development. Interestingly, even Coursera provided partial user justification. The authenticity processes were consistently available in the content review processes. Duolingo and Coursera scored full alignment. Khan Academy's quality assurance processes indicated partial maturity. The data security specifications indicated full implementation for all platforms, reflecting the need for regulatory compliance and being the best-aligned domain for the validation processes. All validation results are included in **Table 3**.

The validation matrix is shown in **Table 3**, where alignment measures show full implementation is focused on data security and human escalation specifications, and partial implementation is predominant in the domains of transparency, tempo governance, boundary governance, and explainability. Khan Academy is demonstrated to have the highest alignment, with six full implementations indicating non-profit nature and development of Khan-

migo, and gaps in tempo governance on all platforms indicate extension of the framework beyond current industry practice.

Table 3. Framework Validation: Platform Evidence Mapping.

Dimension	Component	Duolingo	Coursera	Khan Academy
Governing Dimension				
Governance Principles	Transparency	Partial	Partial	Partial
	User Autonomy	Full	Partial	Full
	Value Alignment	Partial	Partial	Full
Personalization Governance	Intensity	Partial	Partial	Full
	Tempo	Partial	Partial	Partial
	Boundary	Partial	Partial	Partial
Designing Dimension				
VSP System Design	Identity Disclosure	N/A	Partial	Full
	Interaction Norms	N/A	Partial	Partial
	Human Escalation	Partial	Full	Full
AI Promotion System Design	Explainability	Partial	Partial	Partial
	Authenticity	Full	Full	Partial
	Data Security	Full	Full	Full

Note: Full = comprehensive implementation documented; Partial = limited scope or depth of implementation; N/A = element not applicable to current platform configuration.

4. Discussion

The study offers a two-dimensional model that ties user-level trust mechanisms to organizational implementation practices, thus combining two previously disjointed areas of study. This integration enhances the understanding of how trustworthy AI-based marketing systems can be developed in the context of education technology. The validation findings show that full implementation is largely present in regulatory-focused areas such as data security, while it is largely partial in governance-focused areas such as tempo and explainability. There is a pattern here that could be interpreted. The dominance of partial implementations, especially under tempo governance and explainability specifications, suggests two possible interpretations. On one hand, these domains represent actual organizational issues where current industry practice is still behind theoretical needs for trustworthy AI marketing systems. On the other hand, it is possible that the specifications under these domains will require further operationalization to support implementation. In this regard, the validation results provide support for both current relevance and areas for further refinement. Drawing on technology acceptance research in educational settings, it is well established that key UTAUT constructs have substantial predictive ability for behavioral intentions regarding diverse technology uses [26]. These studies focus on the results but rarely examine the governance of the personalization processes within the organization. The three-dimensional model of personalization governance, covering intensity, tempo, and boundary issues, extends the research focus from the results to the principles of operational management, building upon the existing effects.

The framework addresses a gap in technology acceptance research, as past research has focused on individual rather than organizational use of technology. Meta-analytical studies of the UTAUT model demonstrated the predictive validity of its core constructs across various technological settings, while also highlighting the need for extending the model to include implementation processes [26]. Research into the impact of AI marketing interactions has shown that perceived intelligence, anthropomorphism, and quality affect customer outcomes [7]. These user-centered findings are useful but do not provide any indication of how organizations should design and govern this interaction. The designing dimension operationalizes user-centered research findings into actionable technical requirements, directly addressing the gaps identified in the validation, such as VSP identity disclosure and recommendation explainability. Governance scholarship has called for frameworks that integrate ethical principles with operational mechanisms, moving beyond abstractions and toward specifications [27–29]. The two-dimensional approach treats trustworthiness as both an objective for managerial actions and a criterion for design. The validation results show that, on all platforms, there is a preponderance of partial implementation, thereby supporting the need for an integrated approach.

The practical implications clarify the operational challenges faced in educational technology organizations. Studies on large language models in education have emphasized their personalization opportunities, as well as the

need for limitations and mechanisms of governance [29]. The dimension of boundaries is concerned with issues that are not easily operationalized in institutional settings. It does this by specifying data sensitivity classification procedures and functional boundary mechanisms that operationalize policy-level commitments in guidelines for operational implementation. The results of validation show that at present, these guidelines are at the policy level and lack operational specificity. Studies on user trust have identified transparency and explainability as antecedents to user acceptance [30]. Validation indicated partial implementation of the explainability aspect in all the platforms, with no platform having full alignment. This directly corresponds to the gap that the specifications in the framework are meant to resolve.

Additionally, the scope of the framework's application includes the balancing of business interests with the integrity of the educational mission. Research into AI trust indicates that justified trust is not merely based on the performance of the system but also on the ability of the user to evaluate the AI decision-making process [31]. The two-dimensional structure helps to alleviate these impediments by distinguishing the 'governance' aspect from the 'technical execution' aspect. This distinction allows for more systematic allocation of resources. The empirical study of AI-based educational software showed that the design features of the system, such as the quality of interaction and the provision of feedback, have a significant impact on the system's use and learning outcomes, thus highlighting the importance of reliable implementation at the technical level [32]. It is important to note that technical excellence is not necessarily sufficient to guarantee trustworthiness. The validation results, showed that data security had full implementation across all platforms under regulatory compliance pressure, whereas discretionary governance domains showed predominantly partial alignment. It is precisely in these discretionary domains that the framework's systematic guidance is most important for building sustainable relationships between learners and platforms.

The framework has direct implications for policymakers and institutional regulators at all strata of the EdTech ecosystem. At the regulatory level, consumer protection and data privacy regulators in educational settings can utilize the three-dimensional personalization governance model to set standards on the degree of AI marketing communications, tempo alignment specifications that preserve the integrity of learning sessions, and boundary specifications that differentiate between marketing data usage and the prohibited usage of educational behavioral data. The validation finding reveals significantly stronger compliance in regulatory-driven domains, such as data security, than in discretionary domains, such as tempo governance and explainability. This reveals that existing voluntary industry standards have proven insufficient, and that regulatory intervention in governance domains that have been under-implemented might be necessary to improve these areas towards trustworthy operation. At the institutional level, the design dimension specifications can serve as evaluation criteria for EdTech platform partnerships. University procurement committees and school district governance structures could, for example, require documented alignment with VSP identity disclosure and recommendation explainability standards as a requirement for institutional adoption. Moreover, international organizations such as UNESCO, which has established guidelines for human-centric AI in education [19], might consider the framework's organization to be relevant for the design of operational tools for assessment in different regulatory settings.

The limitations of the study should be acknowledged. The validation of the framework was limited by the documentation of the three platforms, which may not be representative of their internal governance practices. The cross-sectional validation also does not account for any changes in governance implementation by the platforms in reaction to gaps identified in their systems. Although the methodological framework incorporates pertinent knowledge domains, additional empirical validation is required. Studies that have investigated organizational implementations and users' perceptions before taking into consideration the prescriptions of the framework are empirically grounded, not theoretically grounded.

Several pathways for future research can be identified based on the results of this study. Longitudinal studies focusing on EdTech organizations that utilize AI marketing systems could clarify the application of the governance frameworks. Survey research focusing on governance could be used to quantitatively test the feasibility of the principles outlined in the framework. Experimental studies of the deployment of the framework can assess its effectiveness in achieving user trust outcomes. Cross-regulatory comparative studies can expand the validation of the framework to platforms operating under different data governance regimes, such as those operating under the Personal Information Protection Law in China or the upcoming AI governance regimes in Southeast Asia. Such research would seek to understand the impact of the regulatory environment on the generality of the framework.

The above validation directions would increasingly transform the dual-dimensional framework from a theoretical construct into an evidence-based tool effective in various environments.

5. Conclusion

This study proposes a dual-dimensional framework for designing and governing trustworthy AI-based marketing systems in educational technology. The research approach is thematic, theory-based, and evidence-based. The concept of trustworthiness is used as a key attribute that links both dimensions of the framework. The framework's alignment with Duolingo, Coursera, and Khan Academy is validated. Regulatory-driven domains have shown acceptable results in implementation. In addition, data security has presented comprehensive alignment across the three platforms. Governance domains have illustrated a different trend compared to the rest. In addition, tempo governance, transparency, and explainability have exhibited partial implementation. Khan Academy has exhibited the best results in terms of alignment. This is because the organization is a non-profit entity and has developed Khanmigo. Duolingo and Coursera have exhibited three comprehensive implementations each. There was a notable absence of alignment in discretionary governance domains, where the framework extends beyond current practice.

The present study contributes to the body of knowledge by filling an established gap in the literature. Previous studies have treated user-level trust mechanisms and organizational implementation as two separate domains. This study proposes a new framework that brings these two areas together in a unified form. The extension of the UTAUT model from the original purpose of predicting individual-level acceptance to the level of organizational-level governance is a theoretically significant shift. It proves the feasibility of systematically translating validated components of individual-level behavior into new forms of value alignment such as user autonomy, transparency, and implementation design standards. The three-dimensional approach to the governance of personalization improves upon other single-dimensional approaches, such as those focusing solely on the concept of intensity, as it also considers tempo alignment with learning cycles and the scope of data usage as separate dimensions of governance, unlike other approaches. The framework is useful for organizations in the educational technology industry, as it offers specifications for the development of AI-based marketing systems. Validation results confirm that tempo governance and the explainability of recommendations are essential for the development of the industry.

Author Contributions

Conceptualization, C.C.; methodology, C.C.; software, C.C.; validation, C.C., N.H.B.Z., W.D., X.X. and Y.X.; formal analysis, C.C., N.H.B.Z., W.D., X.X. and Y.X.; investigation, C.C.; resources, C.C.; data curation, C.C.; writing—original draft preparation, C.C.; writing—review and editing, C.C.; visualization, C.C.; supervision, N.H.B.Z.; project administration, C.C.; funding acquisition, C.C., N.H.B.Z., W.D., X.X. and Y.X. All authors have read and agreed to the published version of the manuscript.

Funding

This work received no external funding.

Institutional Review Board Statement

Not applicable. This study was based entirely on secondary, anonymized, firm-level data made publicly available by the World Bank Enterprise Surveys program, and did not involve any new interaction with human participants or animals.

Informed Consent Statement

Not applicable. The World Bank Enterprise Surveys obtained informed consent from participating firms during the original data collection, and the present study analyzed only de-identified aggregate records.

Data Availability Statement

The firm-level data that support the findings of this study are openly available from the World Bank Enterprise Surveys at https://www.enterprisesurveys.org. Processed datasets and SmartPLS model files are available from

the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflict of interest.

AI Use Statement

The authors used [Claude 3.5 Sonnet (Anthropic, 2024)] solely for grammar checking, sentence structure refinement, and improving the readability of the English text in this manuscript. The authors take full responsibility for all academic content, including all ideas, data, analyses, and conclusions presented herein. The use of AI was thoroughly reviewed and supervised by the authors.

References

1. Huang, M.-H.; Rust, R.T. A strategic framework for artificial intelligence in marketing. *J. Acad. Mark. Sci.* **2021**, *49*, 30–50. [CrossRef]
2. Dwivedi, Y.K.; Hughes, L.; Ismagilova, E.; et al. Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *Int. J. Inf. Manag.* **2021**, *57*, 101994. [CrossRef]
3. Haleem, A.; Javaid, M.; Qadri, M.A.; et al. Artificial intelligence (AI) applications for marketing: A literature-based study. *Int. J. Intell. Netw.* **2022**, *3*, 119–132. [CrossRef]
4. Zawacki-Richter, O.; Marín, V.I.; Bond, M.; et al. Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *Int. J. Educ. Technol. High. Educ.* **2019**, *16*, 39. [CrossRef]
5. Holmes, W.; Tuomi, I. State of the art and practice in AI in education. *Eur. J. Educ.* **2022**, *57*, 542–570. [CrossRef]
6. Gao, Y.; Liu, H. Artificial intelligence-enabled personalization in interactive marketing: A customer journey perspective. *J. Res. Interact. Mark.* **2023**, *17*, 663–680. [CrossRef]
7. Chi, N.T.K.; Hoang Vu, N. Investigating the customer trust in artificial intelligence: The role of anthropomorphism, empathy response, and interaction. *CAAI Trans. Intell. Technol.* **2023**, *8*, 260–273. [CrossRef]
8. Strzelecki, A. To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interact. Learn. Environ.* **2024**, *32*, 5142–5155. [CrossRef]
9. Al-Adwan, A.S.; Li, N.; Al-Adwan, A.; et al. Extending the technology acceptance model (TAM) to Predict University Students' intentions to use metaverse-based learning platforms. *Educ. Inf. Technol.* **2023**, *28*, 15381–15413. [CrossRef]
10. Berendt, B.; Littlejohn, A.; Blakemore, M. AI in education: Learner choice and fundamental rights. *Learn. Media Technol.* **2020**, *45*, 312–324. [CrossRef]
11. Kasneci, E.; Sessler, K.; Küchemann, S.; et al. ChatGPT for good? On opportunities and challenges of large language models for education. *Learn. Individ. Differ.* **2023**, *103*, 102274. [CrossRef]
12. Gołąb-Andrzejak, E. AI-powered digital transformation: Tools, benefits and challenges for marketers–case study of LPP. *Procedia Comput. Sci.* **2023**, *219*, 397–404. [CrossRef]
13. Venkatesh, V.; Thong, J.Y.; Xu, X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Q.* **2012**, *36*, 157–178. [CrossRef]
14. Bag, S.; Srivastava, G.; Bashir, M.M.A.; et al. Journey of customers in this digital era: Understanding the role of artificial intelligence technologies in user engagement and conversion. *Benchmarking* **2022**, *29*, 2074–2098. [CrossRef]
15. Shin, D. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *Int. J. Hum.-Comput. Stud.* **2021**, *146*, 102551. [CrossRef]
16. Ferrario, A.; Loi, M. How explainability contributes to trust in AI. In Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency, Seoul, Republic of Korea, 21–24 June 2022; pp. 1457–1466. [CrossRef]
17. Chan, C.K.Y. A comprehensive AI policy education framework for university teaching and learning. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 38. [CrossRef]
18. Chiu, T.K. The impact of Generative AI (GenAI) on practices, policies and research direction in education: A case of ChatGPT and Midjourney. *Interact. Learn. Environ.* **2024**, *32*, 6187–6203. [CrossRef]
19. Miao, F.; Holmes, W. *AI and Education: Guidance for Policy-Makers*; UNESCO Publishing: Paris, France, 2021.

20. Okonkwo, C.W.; Ade-Ibijola, A. Chatbots applications in education: A systematic review. *Comput. Educ. Artif. Intell.* **2021**, *2*, 100033. [CrossRef]
21. Chen, Y.; Jensen, S.; Albert, L.J.; et al. Artificial intelligence (AI) student assistants in the classroom: Designing chatbots to support student success. *Inf. Syst Front.* **2023**, *25*, 161–182. [CrossRef]
22. Essel, H.B.; Vlachopoulos, D.; Tachie-Menson, A.; et al. The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *Int. J. Educ. Technol. High. Educ.* **2022**, *19*, 57. [CrossRef]
23. Jabareen, Y. Building a conceptual framework: Philosophy, definitions, and procedure. *Int. J. Qual. Methods* **2009**, *8*, 49–62. [CrossRef]
24. Braun, V.; Clarke, V. *Thematic Analysis: A Practical Guide*; Sage Publications Ltd: Los Angeles, LA, USA, 2021.
25. Xue, L.; Rashid, A.M.; Ouyang, S. The unified theory of acceptance and use of technology (UTAUT) in higher education: A systematic review. *SAGE Open* **2024**, *14*, 21582440241229570. [CrossRef]
26. Tamilmani, K.; Rana, N.P.; Dwivedi, Y.K. Consumer acceptance and use of information technology: A meta-analytic evaluation of UTAUT2. *Inf. Syst Front.* **2021**, *23*, 987–1005. [CrossRef]
27. García de Blanes Sebastián, M.; Sarmiento Guede, J.R.; Antonovica, A. Application and extension of the UTAUT2 model for determining behavioral intention factors in use of the artificial intelligence virtual assistants. *Front. Psychol.* **2022**, *13*, 993935. [CrossRef]
28. Filgueiras, F. Artificial intelligence and education governance. *Educ. Citizensh. Soc. Justice* **2024**, *19*, 349–361. [CrossRef]
29. Labadze, L.; Grigolia, M.; Machaidze, L. Role of AI chatbots in education: Systematic literature review. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 56. [CrossRef]
30. Leichtmann, B.; Humer, C.; Hinterreiter, A.; et al. Effects of Explainable Artificial Intelligence on trust and human behavior in a high-risk decision task. *Comput. Hum. Behav.* **2023**, *139*, 107539. [CrossRef]
31. de Brito Duarte, R.; Correia, F.; Arriaga, P.; et al. AI trust: Can explainable AI enhance warranted trust? *Hum. Behav. Emerg. Technol.* **2023**, *2023*, 4637678. [CrossRef]
32. Chang, C.Y.; Hwang, G.J.; Gau, M.L. Promoting students' learning achievement and self-efficacy: A mobile chatbot approach for nursing training. *Br. J. Educ. Technol.* **2022**, *53*, 171–188. [CrossRef]

Copyright © 2026 by the author(s). Published by UK Scientific Publishing Limited. This is an open access article under the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Publisher's Note: The views, opinions, and information presented in all publications are the sole responsibility of the respective authors and contributors, and do not necessarily reflect the views of UK Scientific Publishing Limited and/or its editors. UK Scientific Publishing Limited and/or its editors hereby disclaim any liability for any harm or damage to individuals or property arising from the implementation of ideas, methods, instructions, or products mentioned in the content.