ATU PRESS

ISSUES IN LANGUAGE TEACHING (ILT)
VOL. 14, NO. 1, 163-212, June 2025
https://doi.org/10.22054/ilt.2025.88721.939 Document Type: Research Paper

A Contextualized Co-Design Framework for Integrating Generative AI into Iranian ESP Education

Masoud Azadnia
Assistant Professor of TEFL, Islamic Azad University, Naein, Iran.

Received: March 08, 2025**; Revised:** June 20, 2025**; Accepted:** June 25, 2025

Abstract
Although generative artificial intelligence (GenAI) is reshaping language education globally, its integration into English for Specific Purposes (ESP) instruction in Iran has remained limited by rigid curricula, restricted autonomy, and a lack of authentic, discipline-specific resources. This study sought to address this gap by developing a contextualized co-design framework to align AI affordances with local pedagogical realities. Guided by theoretical triangulation, a hybrid methodology was employed that combined design-based research, co-design, and participatory qualitative approaches. The needs analysis phase involved 26 ESP stakeholders, including eight instructors and 18 learners. In-depth interviews were then conducted with a subset of 19 participants—eight instructors and 11 learners. The final co-design phase engaged a diverse group of stakeholders, including 12 instructors, learners, policymakers, and AI experts, to incorporate multiple perspectives into the model development process. As revealed by the findings, the learners prioritized personalization, writing support, and disciplinary adaptability, whereas the instructors emphasized controlled prompting, ethical literacy, and assessment redesign. Concerns shared between the two groups included overreliance, epistemic authority, and unequal digital access. The co-design process generated a cyclical instructional model incorporating dual human-check mechanisms, ethical self-reporting, bias-awareness checkpoints, multimodal feedback loops, and institutionalized teacher training. Offering a theoretically grounded and adaptable reference for discipline-specific AI use in ESP programs, this study's framework can guide educators, curriculum designers, and policymakers in ESP contexts analogous to those investigated in this research.

Keywords: Co-design framework, Design-based research, English for Specific Purposes (ESP), Generative Artificial Intelligence (GenAI), Iranian higher education

Author's email: masoudazadnia@gmail.com

INTRODUCTION

Technological affordances of generative artificial intelligence (GenAI) tools, the state-of-the-art technology of the time, have evolved the routine practices in many fields of education, particularly English language learning and teaching, by promoting personalized learning, tailored discipline-specific content, and active learner engagement (Weng & Fu, 2025). The growing body of research over the past couple of years since the release of these smart tools (e.g., Abusahyon et al., 2023; Jegede, 2024; Krishnan & Zaini, 2025; Nguyen, 2025; Pan, 2024; Shakibaei et al., in press; Wu, 2024) has demonstrated their contributory role in enhancing various English learning outcomes, including writing, vocabulary development, and academic discourse rehearsal. The recent research line also accentuates the positive impact of this interactive, adaptive, and game-like learning atmosphere on learner motivation and self-directedness. The undeniable virtues of GenAI tools notwithstanding, the research evidence offers mixed conclusions about the most effective ways of embedding such technologies into mainstream language education.

Serving as an indispensable part of the Iranian higher education curriculum, English for Specific Purposes (ESP) occupies a space between general language proficiency and discipline-specific literacy, equipping students with the communicative skills needed for scholarly work (Atai & Tahririan, 2003). Yet, even with ESP's central place in the curriculum, progress is often stymied by inflexible syllabuses, minimal input from learners, limited subject matter expertise of instructors, and a scarcity of real-world resources (Eslami, 2010; Mostafavi et al., 2021; Rajabi et al., 2012). GenAI tools influence ESP significantly by tailoring instruction to the specific professional and academic communication needs of learners (Elmotri et al., 2025). Although these advanced technological tools have shown great promise in ESP by enabling automatically tailored content, highly immersive and interactive simulations, and adaptive teaching methods that respond to individual differences, they also demand careful and sustained management to prevent unintended consequences. Luckin et al. (2016) caution that, in the absence of disciplined, careful exploration, such innovations tend to intensify, rather than close, the enduring distance between curricular intent and the affordances of the digital tool, generating fragmentation in syllabus coherence, ethical governance, and infrastructural inequality across contexts.

Introducing these systems to environments outside the global North, Iran included, falls far beyond a straightforward engineering undertaking; it also requires addressing divergent digital literacies, patchy institutional readiness, and deeply interwoven socio-cultural codes that collectively mediate educational practice (Derakhshan et al., 2025). The aggregation of these factors obliges a primary and unequivocal imperative: the conception and rollout of ESP pedagogies enabled by artificial intelligence (AI) must be inextricably anchored to the prevailing regional educational topography, must be regulated by ethical conventions whose rationale and applicability can be unequivocally defended, and must be calibrated to the operational and human terrains of the educational systems and their constituents.

Inspired by the research-driven priority of deliberate, contextually attuned integration, this study aims to develop a hybrid model for generative AI in ESP teaching in Iran, grounded in theoretical and methodological triangulation and incorporating elements of interpretative, design-based, and co-design approaches. The triangulated theoretical framework was based on the Technological Pedagogical Content Knowledge (TPACK) model (Mishra & Koehler, 2006), the sociocultural theory (SCT; Lantolf & Thorne, 2006; Vygotsky, 1978), and the human-centered AI principles. Based on SCT, the scaffolding and mediating role of GenAI tools directs learners' interactions, helping them progress within their Zone of Proximal Development (ZPD). The TPACK model calls for striking a balance among technology, pedagogy, and content knowledge, while reflecting on how to integrate AIs in mainstream educational settings, such as conventional ESP contexts, where instructors are expected to possess a broad array of intellectual assets (Maghsoudi, 2023). In line with these frameworks, human-centered AI stresses that ethical principles, cultural sensitivity, and equity should guide the adoption of GenAI tools. The triangulated framework suggests that AI in ESP education should not be viewed as a mere technical tool. Instead, it represents a pedagogical practice shaped by interaction, teacher expertise, and ethical responsibility.

LITERATURE REVIEW

GenAI and ESP: Opportunities and Pitfalls

The rapid roll-out of GenAI tools in the specific realm of ESP, where high-level linguistic performance is characterized by subject-specific accuracy and adherence to professional standards, has sparked a wide spectrum of reactions varying from unbounded enthusiasm to profound apprehension. Early evidential data on AI-assisted language pedagogy (e.g., Ahmed et al., 2025; Lai, 2025) have testified several favorable outcomes, such as drafting acceleration, writing burden reduction, cognitive load reduction, and structured lexical repertoire expansion. Appealing and rewarding notwithstanding, these attainments are difficult to generalize to deeper disciplinary competence, since AI models often prioritize overall coherence over domain-specific accuracy. This priority accentuates the model’s weakness in supporting advanced ESP tasks highly contingent on lexical precision and discourse convention adherence (Demirdöken, 2024; Puspasari & Agustina, 2025; Zhyhadlo & Zaiarna, 2025).

A detailed scrutiny of the research body on the educational consequences of GenAI tools shows that promising educational outcomes are mainly the natural product of integrating these tools into structured pedagogical practices, such as peer- and teacher-facilitated feedback, guided revisions based on rubrics, or sequential drafting stages (Giannakos et al., 2024; Holstein et al., 2019; Noroozi et al., 2024; Renfeng et al., 2025). On the contrary, GenAI facilities used in isolation, detached from scaffolded pedagogical routines, inevitably spark automation-induced bias and overreliance, two major technology use concerns that reduce autonomy and reflective language awareness (Artyukhov et al., 2025; Marchena Sekli et al., 2024; Zaim et al., 2025). In ESP settings, where professional standards and disciplinary credibility are highly consequential, such risks are particularly pronounced. The literature therefore suggests that effective integration requires clarity on the division of labor between human actors and AI systems, yet most current studies under-specify these human "gatekeeping" points.

Constraints on Adopting GenAI tools in ESP Pedagogy

Current literature acknowledges that assessment frameworks fundamentally influence the efficacy of GenAI in language learning. Predominantly product-oriented evaluation systems, prevalent especially in ESP programs under centralized governance, allow learners to submit polished AI-generated texts that may lack authentic disciplinary mastery, a phenomenon termed the "rebound effect" (Francis et al., 2025; Smith et al., 2025; Zawacki-Richter et al., 2019). On the other hand, as implied by empirical data (e.g., Weng et al., 2024; Xia et al., 2024), process-oriented assessments, such as revision monitoring and oral examinations/defenses, not only can establish assessment validity but also reduce blind dependence on AI output.

Furthermore, equity challenges, such as hardware availability disparities, digital literacy differences, and instructor professional development variations, pose significant barriers to widespread adoption of AI-enhanced teaching and learning (Jia, 2025; Pawar& Khose, 2024). The ways through which equity issues enumerated above can hinder AI integration in environments lacking strong instructional and infrastructural support are elucidated in a couple of recent studies (e.g., Dolba et al., 2023; Zipf et al., 2025).

Foresight in continuous professional development is another essential tenet; without it, AI integration in ESP education risks remaining fragmented and short-lived, undermining instructional quality and reinforcing unequal access to its benefits. The absence of sustained professional development programs, as supported by evidential data (e.g., Roshan et al., 2024; Shrestha et al., 2025), may be attributed to the TPACK framework's flaw in capturing the specialized expertise required to align AI tools with the complex needs of ESP disciplines.

Contextual Challenges in the Iranian ESP Context

The Iranian ESP context faces a range of instructional challenges that affect both teaching practices and learning experiences. One of these challenges is the lack of ongoing development in professional knowledge and digital literacy aligned with emerging technologies among ESP instructors. This limitation, as evidenced by research (e.g., Dashtestani & Stojković, 2016; Ghiasvand et al., 2024; Nezakatgoo & Behzadpoor, 2017), may hinder meaningful and effective implementation of AI-enhanced pedagogies. Additionally, as supported by evidential data (e.g., Mostafaei Alaei, 2016; Mostafavi et al., 2021), rigid, high-stakes assessments and prescribed teaching goals embedded in ESP curricula provide little room to develop AI-driven, discipline-specific discourse, a core demand of the future workplace. Such pedagogical constraints do not exist in isolation; they interact with persistent infrastructural fragility. Day-to-day access to learning AI tools is still disrupted by erratic connectivity, aging computers, and unpredictable cycles of basic tech support (Hosseini Moghadam, 2023). Taken together, these administrative and hardware characteristics compound the uneven ability of schools to integrate AI effectively, marginalizing learners for whom consistent, quality, and prepared exposure to technology is most essential to the pedagogical goals they otherwise pursue (Parviz, 2024).

Research Gap and Questions

In summary, the existing literature reflects four critical gaps: insufficient validation of GenAI's applicability to discipline-specific ESP communication; unclear delineation of human and AI roles in pedagogy; predominance of product-based rather than process-sensitive assessment; and scarcity of context-sensitive, longitudinal, co-designed models, particularly within Iranian educational realities. This study addresses these deficits by collaboratively exploring the perceptions, needs, and constraints of Iranian ESP learners and instructors engaged with AI-assisted ESP instruction, as well as co-designing a contextualized, sustainable framework for GenAI integration. The research questions are as follows:

1. What perceptions, needs, and contextual challenges do ESP instructors and ESP learners hold regarding the integration of generative AI into Iranian ESP education?
2. How can a co-design process involving ESP instructors, ESP learners, university policymakers, and AI specialists produce a contextualized model for using generative AI tools in Iranian ESP classrooms?

METHOD

Research Design

The current study benefited from a combined research design that integrated the principles of a qualitative interpretative approach (Adam, 2020), co-design methodology (Sanders & Stappers, 2014; Spinuzzi, 2005), and design-based research (DBR) (McKenney & Reeves, 2012) to fulfill its specific purposes. This integrated methodology was informed by a theoretical triangulation, which employed the tenets of the SCT, the TPACK framework, and human-centered AI principles in designing data collection instruments, developing coding schemes, and interpreting the findings. SCT informed the exploration of contextual and interactional factors in AI-assisted ESP instruction, while the TPACK framework directed attention to the intersection of technological, pedagogical, and content knowledge. On the other hand, the research's focus on stakeholder-centered design and ethical considerations, as manifested through ethical concern cards and scenario tasks in the co-design phase, was grounded in the tenets of the human-centered AI theory. This theoretical triangulation ensured the comprehensiveness of data collection and analysis procedures, capturing all cognitive, technological, pedagogical, and ethical dimensions in the development of the target model.

The use of an interpretative design in the initial stage of the work suited the core objective of the study: to explore the pedagogical and learning needs, perceptions, and contextual challenges of key stakeholders involved in AI-assisted ESP instruction in an Iranian EFL context. A purposive sample of ESP instructors and learners was asked to share their lived experiences and the socio-cultural factors shaping their engagement with AI in ESP instruction. The co-design approach, on the other hand, facilitated the development of a contextualized pedagogical model for AI integration in ESP instructional routines. While co-design originated primarily in digital tool development, research (e.g., Holmlid, 2009; Luckin et al., 2016) confirms its practical applicability in developing pedagogical frameworks. The DBR tenets called for iterative modeling and active, real-world engagement of the stakeholders in developing a context-responsive pedagogical model suited to contextual realities, ethical considerations, and stakeholder needs.

A three-phase data collection process was utilized to (1) analyze the context-specific needs, perceptions, and challenges pertinent to an AI-assisted ESP instruction; (2) deepen and validate the data drawn from the needs analysis and contextual exploration phase through semi-structured interviews; and (3) finalize and verify the model through co-design participatory workshops. The generative toolkit used in the co-design phase was developed by jointly analyzing individual data from the first and second phases. In the third phase, the outputs of each workshop were iteratively analyzed as inputs in future ones, thereby allowing continuous refinement and validation of the emergent model. (see Figure 1).

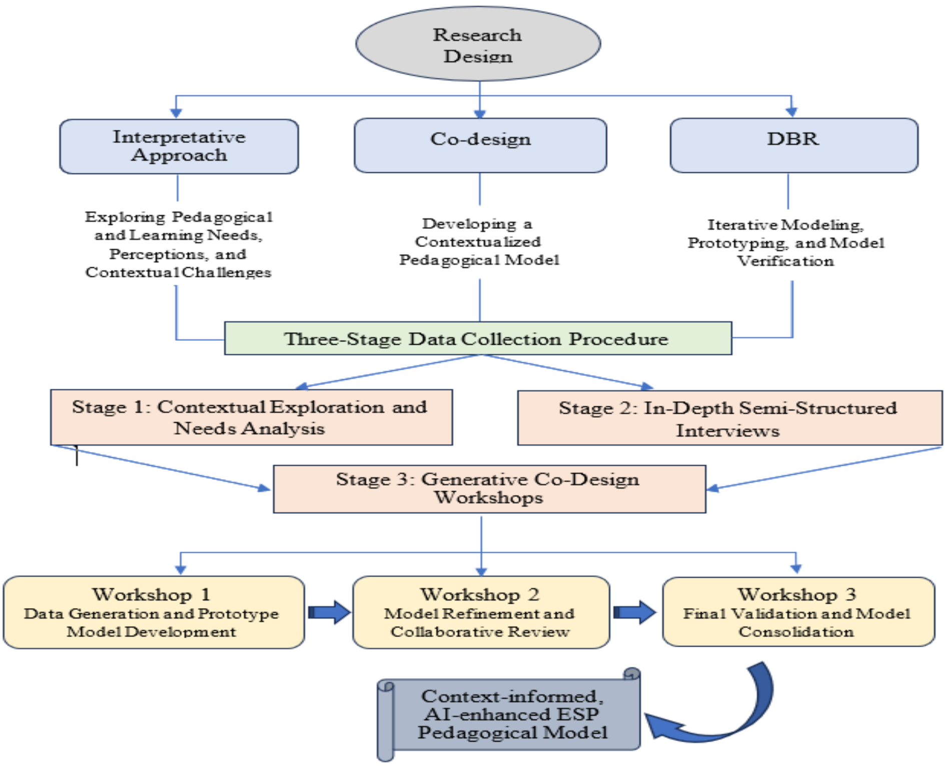


Figure 1. *Research Design and Procedure*

Participants

A 26-member, purposively chosen sample of ESP teachers and learners, as the target models' stakeholders, was recruited to conduct the introductory stage of the study. The participants included eight ESP instructors and 18 ESP learners from four Islamic Azad University branches in Iran, including Isfahan (Khorasgan), Najafabad, Shahreza, and Naein. The author's prior institutional affiliation, which enhanced the logistical feasibility of a sustained field engagement throughout the study, was the core reason for delimiting the participant population to university branches enumerated above. The key criteria for the instructor sample included five years (or more) of ESP teaching experience, prior experience integrating an AI tool into ESL teaching practices, and willingness to participate in the study. To select the learner participants, snowball sampling was employed, and the instructors were asked to identify learners within their classrooms who have first-hand experience with using one or more forms of AI in ESP learning and are willing to participate in the study. The demographic information for Phase I participants is presented in Table 1.

Table 1. *Demographics of Phase I Participants*

Group	N	Gender		Age (M)	Academic Background
		Male	Female		
Instructor	8	5	3	52.6	Teaching English as a Foreign Language (TEFL), English Language Translation, Industrial Engineering, Microbiology, Computer Engineering
Learner	18	7	11	22.8	(Automotive, Mechanical, Computer, Electronic, Civil, Industrial) Engineering, English Language Translation, Psychology, Urban Planning, Microbiology, Chemistry, Physical Education, Nursing, Architecture, Accounting, Educational Science, Banking Management, Media Management, Curriculum Planning

From among those who participated in Phase I, a 19-member group, containing all eight ESP instructors and 11 out of the 18 ESP learners, expressed their availability and willingness to proceed with the study, attending in-depth semi-structured interviews. The selection of individuals who had previously participated in phase I ensured the depth and continuity of data collection, thereby allowing for triangulated, context-informed qualitative data collection.

Of all the participants from prior phases, three ESP instructors and three ESP learners were invited to participate in the co-design phase. Along with these six key stakeholders, three university-affiliated policymakers (two language department chairs and one university president) and three AI experts (two specialists in educational technology and one human-centered AI design expert) participated in the co-design phase. The inclusion of different groups of stakeholders enhanced model comprehensiveness, ensuring due regard for various learning, teaching, executive, and technological considerations. The co-design team members were purposively selected based on their active engagement in earlier phases (as for ESP instructors or learners), disciplinary relevance, and willingness to participate in the co-design stage.

The participants' involvement in every research phase required their written consents and full awareness of the objectives and procedures. They were also assured of anonymity and secure storage of their data. In addition to adhering to ethical research standards, such as those outlined by the Institutional Review Board (IRB), the study acknowledged AI-related ethical challenges, particularly during the co-design phase when developing ethical concern cards to encourage active participation and engagement. The participants were assured of their right to withdraw at any stage.

Instruments

Different qualitative data gathering instruments, suited to participatory and co-design tenets (Adam, 2020; Sanders & Stappers, 2014; Spinuzzi, 2005), were employed to address the research questions. The instruments were all pilot tested and administered based on pre-determined protocols to enhance instrument reliability and implementation rigor, respectively.

Cultural Probes (CPs) and Participatory Mapping Tasks (PMTs)

CPs were used to explore the participants' perceptions of and experiences with AI use in ESP contexts on an individual basis. They were structured around three main parts, including (a) a daily reflection card, which probed into learners' and instructors' perceptions about effective pedagogical tasks, use of AI tools, and challenges faced in their ESP classes; (b) a class day mapping sheet, which asked the respondents to mark their class structure, AI used in it, as well as confusion/success points found; and (c) a speculative AI prompt, which asked the participants to imagine and describe the affordances, consequences, and challenges of an AI assistant in their ESP class. Designed in two different versions for ESP instructors and learners, the instrument consisted of an instructor version that probed into lesson success, AI tool usage, and instructional adaptation, whereas the learner version focused on learning supports, AI experiences, and points of confusion.

The other instrument used for data gathering, PMTs, included two task categories: scenario sketching and ESP environment mapping. The first task involved both ESP instructors and learners plotting either a helpful (ideal) or an unhelpful (problematic) AI-enhanced ESP class, to determine the distinct roles of all stakeholders (AI, instructor, and learner). The second task asked them to draw their own ESP class (physical/digital) spaces by highlighting supports, hindrances, and communication processes thereof. The tasks were identical for both instructor and learner participants.

Semi-Structured Interview Protocol

Semi-structured interviews were used to explore the practices and perceptions of ESP instructors and learners regarding the integration of AI in ESP pedagogy. Every interview entailed answering 12 open-ended questions about participant background and digital experiences (two questions), AI conceptualization and practice (two questions), AI opportunities and challenges (two questions), instructional roles and ethical principles of AI use (three questions), and AI design/trust factors (three questions). In alignment with the triangulated theoretical framework of the study (socio-cultural, TPACK, and human-centered AI theories), the interview questions covered cognitive, pedagogical, technological, and ethical dimensions to provide an in-depth account of AI-assisted ESP instruction. The interview protocol entailed using concise prompts to elicit detailed and in-depth answers.

Generative Co-design Toolkit

A generative co-design toolkit, including (a) role insight cards (RICs), (b) ethical concern cards (ECCs), (c) scenario storyboard templates (SSTs), and d) a set of pre-printed modeling stickers and icon sheets for flowchart construction, was designed to gather participatory-generated co-design data. The toolkit's distinct parts were based on the core themes that emerged from analyzing the key stakeholders' learning/pedagogical routines, experiences, and expectations. Synthesizing the participants' pedagogical and learning profiles from their responses to PMTs and semi-structured interviews, the analysis enabled perspective-based role adoption via RICs throughout the co-design tasks. ECCs promoted context-situated reflection planning, presenting primary dilemmas, tensions, and ethical concerns regarding human-AI interactions drawn from sub-themes that emerged from phase I and II findings. SSTs were intended to integrate ethical and pedagogical insights through blank narrative frameworks, asking different participant groups to depict practical scenarios of optimal and problematic AI-enhanced ESP instruction. The pre-designed, printed modeling stickers and icon sheets included standardized shapes (rectangles for actions, diamonds for decisions, circles for start/end) and pictograms for roles (teacher, learner, AI, etc.) and functions (feedback, risks, checkpoints, etc.).

Data Collection Procedure

The data collection process was structured into three sequential and iterative phases. Table 2 provides a concise overview of the purpose, main activities, data sources, and participants for each phase.

Table 2. *Overview of the Three-Phase Data Collection Procedure*

Phase	Purpose	Main Activities	Data Sources	Participants
Phase I	Contextual exploration & needs analysis	Cultural Probes (CPs), Participatory Mapping Tasks (PMTs), workshop-style administration	Reflection cards, class mappings, speculative prompts, sketches	8 ESP instructors + 18 ESP learners (N=26)
Phase II	Deepening and triangulation of Phase I findings	In-depth semi-structured interviews	Interview transcripts	8 ESP instructors + 11 ESP learners (N=19)
Phase III	Participatory model development & validation	Three iterative co-design workshops using generative toolkit (RICs, ECCs, SSTs, flowcharts)	Role/ethical cards, storyboards, flowcharts, audio recordings, researcher notes	3 ESP learners + 3 ESP instructors + 3 policymakers + 3 AI experts (N=12)

Phase I: Contextual Exploration and Needs Analysis

The initial data gathering phase, which entailed eliciting data from both participant groups' pedagogical/learning experiences, challenges, and expectations regarding AI integration into ESP instruction, unfolded across two stages. During the first phase, all participants received face-to-face guidance from the researcher on responding to the CPs, which could be completed either in person after class or independently at home and submitted before the following session. Over the data collection span, the researcher remained available to answer questions and provide both online and in-person assistance. The second stage entailed eliciting data through PMTs administered in participatory workshops at the four university branches. To this end, during a 90-minute workshop, the ESP instructors and learners participated in group discussions, narrative elaboration, and sketching activities required by PMTs. The workshop sessions were held in Persian to enhance cultural and linguistic relevance.

Phase II: In-depth Semi-structured Interviews

The second research phase entailed in-depth semi-structured interviews with ESP learners and instructors regarding their learning, teaching, ethical, and conceptual ideas pertinent to an AI-assisted ESP pedagogy. The interview consisted of 12 open-ended questions, categorized under six probing themes (two questions per theme), including digital background, conceptual understanding of AI use, opportunities and challenges of AI use, learning and teaching roles, ethical and trust concerns, and design expectations. Based on the participants' convenience and choice, the interviews, each lasting between 30 and 40 minutes, were conducted either face-to-face or virtually. The preplanned interview protocol encouraged a conversational style in interviewing and asking follow-up clarification questions when needed to guarantee data authenticity and clarity. Though all interviews were conducted and transcribed in Persian, the selected transcripts for thematic analysis and report were rendered into English.

Phase III: Generative Co-Design Workshops

The co-design phase comprised three participatory workshops held over two weeks at a private venue arranged by the researcher. All three workshops involved the participation of the whole 12-member co-design team.

Conducted in Persian, the workshops prioritized clear communication and cultural relevance. The procedure for generative data gathering is detailed as follows:

Workshop 1

The first workshop marked the progression from exploratory data collection (Phases I & II) toward the participatory and generative design. It was a five-hour participatory session with a 30-minute break to maintain productivity that entailed going through a protocol aligned with the tasks and activities in the generative co-design toolkit. First, the co-design team was split into three groups with members of mixed roles to encourage balanced dialogue, shared sense-making, and interdisciplinary collaboration. Then, using RICs, every group was involved in jointly mapping role-specific needs, values, and concerns. Next, every mixed-role group used ECCs to examine short speculative cases, highlighting common ethical issues in AI-enhanced ESP (e.g., bias, overreliance, trust) and document proposed resolutions and underlying reasoning on annotated sheets. Subsequently, the members in each group drafted their ideal and problematic instructional scenarios individually, integrating role priorities, ethical considerations, and process models. The initial sketches of scenarios were refined through participatory group discussions. Lastly, every group embarked on co-creating prototype instructional flows, defining communication pathways, feedback loops, and task sequences using layout templates and pre-determined role and action icons/shapes. Given the workshop's in-person setting without digital collaboration tools, the co-design team employed physical materials, such as large poster boards and printed paper-based stickers/icons, to represent roles and actions.

During the first generative workshop, the researcher took a dual facilitator-observer role, whereby he confined his intervening role to providing highly required guidance and clarification explanations. Based on his observer role, he adopted a protocol that included turn-taking, role rotation, and lost voice evocation strategies to promote equitable participation, open conversation, and respectful listening during the workshop. All these minimal interventions were documented in observational memos to ensure transparency and reflexivity.

Workshop 2

Having analyzed the data from workshop one, a second workshop of three hours was held to review and refine the emerging topics and preliminary models obtained from the initially gathered data. By sharing their experiences and implementing specific actions, the participants offered feedback and recommendations to enhance the AI-assisted ESP teaching methodology. In this stage, unlike the group-based structure of Workshop 1, all twelve participants engaged in whole-group discussions where each member had the chance to review, question, and refine the synthesized results. Divergences across the prototype flow diagrams were collectively examined until consensus was reached, ambiguities were clarified by the researcher or peers, and proposed expansion areas were added whenever supported by a majority agreement. The researcher guided the process by providing compiled findings, facilitating dialogues among participants, and taking observational notes.

Workshop 3

Following another week of analysis, the final two-hour session was conducted to validate and finalize the pedagogical model with participant consensus. In the last workshop, all team members had their individual turn to openly express their final views and considerations after collaboratively reviewing the refined model developed based on the refinements proposed in Workshop 2. The participants were asked to have due regard for evaluating the models' trustworthiness and applicability in the target context while working in their groups. The researcher's role in this workshop was confined to observing group-based, consensus-making practices and documenting individually articulated, final feedback and reflections.

Data Analysis

The thematic schemes from analyzing Phases I and II data were integrated and triangulated to develop a comprehensive response to the first research question, which explored the needs, perceptions, and contextual challenges of key stakeholders regarding an AI-assisted ESP instruction. Phase I data encompassed a multifaceted collection of qualitative inputs, including daily reflection cards, class mapping sheets, artifact pictures, prompted narrative transcriptions, field notes, audio-video recordings, scenario sketches, environment diagrams, and different visual artifacts. Semi-structured interview transcripts constituted the Phase II data. No predefined template guided the thematic analysis, thereby allowing the emergence of inductive data-driven themes suited to the study's design-based goal of developing a context-relevant, AI-assisted ESP instruction model through iterative co-design. Once emerged, the inductive themes were categorized under the three overarching inquiry themes probed by the first research question.

Regarding the second research question, the analytical focus centered primarily on the co-design phase. The qualitative dataset generated in the workshops, including textual data gathered through IRCs and ECCs, audiovisual recordings of the sessions, prototype instructional flowcharts, and scenario storyboards, was subjected to thematic analysis to define meaningful connections between core emergent themes during phases I and II, including pedagogical routines and practices, ethical considerations, participant roles, and AI integration affordances. The findings from every workshop informed the tasks and procedures in the next one to encourage team alignment, iterative collaboration, and continuous model improvement. By the end of each workshop, the outcomes were scrutinized again to identify any necessary modifications and obtain final approval from the co-design team.

The analysis of textual data gathered across all phases followed a consistent, inductive framework based on Braun and Clark's (2006) six-step thematic analysis rubric, including familiarization, initial coding, initial theme generation, theme review, final theme generation, and report development. The analytical approach aligned with the interpretive and design-based nature of the study, providing a rich, context-specific thematic scheme that emerges naturally from diverse data sources without reliance on predefined coding templates. The entire analytical process was conducted manually to capture subtle, context-specific meanings in participants' responses.

The analysis incorporated visual and material artifacts alongside textual data through an artifact-based approach (Bagnoli, 2009; Rose, 2016). The Phase I visual data consisted of reflection cards and scenario sketches from cultural probes, whereas Phase III generated storyboards and prototype instructional flowcharts. These visual data underwent a two-tier analytical process: first documenting visual elements like symbols and annotations, then interpreting their functional purpose by identifying AI decision points, human review stages, and risk indicators. These visual findings were systematically cross-referenced with textual themes to ensure comprehensive interpretation of stakeholder perspectives.

Intercoder reliability was checked to ensure the analytic rigor and to counter any potential bias. Having conducted the whole analytical process, the researcher asked an independent expert in ESP instruction, familiar with the research objectives, to review samples of coded transcripts and visual materials. The intercoder reliability index (Cohen's $\kappa = 0.79$) proved acceptable; however, all coding inconsistencies were reviewed and rectified through consensus conversations.

RESULTS

The results are presented in two main sections that directly correspond to the two research questions, thereby enhancing readability and structural clarity. The first section synthesizes ESP instructors' and learners' perceptions, needs, and contextual challenges related to the integration of generative AI in Iranian ESP education (RQ1). The second section describes the participatory co-design process and presents the resulting contextualized pedagogical model (RQ2).

Addressing the First Research Question

Results From CPs and PMTs

Tables A1 and A2 in the Appendix detail the thematic analysis results related to the ESP instructors and learners, respectively. As seen in these tables, the themes and subthemes that emerged from the instructor- and learner-elicited data differed to a great extent. There were points of convergence regarding challenges in AI use, such as concerns about overreliance on AI output and access inequities, but the emergent themes diverged in most cases. A clear example is the instructors' emphasis on controlled prompting and assessment reframing, which contrasts with the learners' focus on personalized support and discipline-specific adjustments. These salient discrepancies in themes and subthemes underlined the necessity of an independent analysis of the thematic results drawn from ESP instructors and learners. Given that the research question explored the two groups' needs, perceptions, and contextual challenges with respect to AI-enhanced ESP instruction, all emergent themes and subthemes were subsequently grouped under three analytical categories: Pedagogical/Learning Needs, AI Integration Perceptions, and Contextual Challenges/Concerns (see Table 3).

As Table 3 shows, with respect to Needs, the ESP instructors emphasized pedagogical adaptations such as controlled prompting, reframed assessments, and fostering ethical awareness. The ESP learners, by contrast, voiced the need for personalized support, including AI-assisted vocabulary and writing development, language proficiency scaffolding, and field-specific adjustments, while also calling for a teacher-led human–AI balance. Regarding AI integration, both ESP instructors and learners warned of the overreliance hazard; the ESP instructors further pointed to reduced autonomy and questioned AI's accuracy in specialized subjects, whereas the ESP learners, though more hopeful and reporting positive experiences, also raised concerns over fairness and academic honesty. Finally, under Contextual Challenges, the ESP instructors underscored sequencing and prioritization difficulties, while both groups cited digital literacy gaps and unequal access.

The ESP learners additionally reported frustrations with AI misinterpretations and limitations tied to institutional constraints.

Table 3. *Thematic Analysis Results based on the Data from CPs and PMTs*

Category	Theme	Subtheme	Stakeholder
Pedagogical /Learning Needs	Learning Supports	AI-based Vocabulary Aids	Learner
	Pedagogical Adaptations	Controlled Prompting Strategies	Instructor
	Pedagogical Adaptations	Reframing Assessment Tasks	Instructor
	Pedagogical Adaptations	Ethical Awareness Development	Instructor
	Learning Supports	Writing Assistance	Learner
	Expectations/Attitudes	Human-AI Balance	Learner
	Learning Supports	Personalized /Customized Learning	Learner
	Learning Context Influences	Language Proficiency Supports	Learner
	Learning Context Influences	Discipline-Specific Needs	Learner
AI Integration Perceptions	Expectations/Attitudes	Optimism about AI Potential	Learner
	AI Integration Experiences	Positive Engagement	Learner
	AI Use Challenges	Overreliance/Reduced Autonomy	Instructor
	Expectations/Attitudes	Ethical and Academic Concerns	Learner
	AI Use Challenges	Doubt in Epistemic Authority	Instructor
	AI Use Challenges	Overreliance on AI	Learner
Contextual Challenges /Concerns	AI Use Challenges	Digital Literacy Gaps	Learner
	AI Integration Experiences	Access Inequity	Learner
	AI Integration Experiences	Confusions/Frustrations	Learner
	AI Use Challenges	Unequal Access & Digital Literacy	Instructor
	Pedagogical Routines	AI Use in Routine Instructions	Instructor
	Pedagogical Routines	Task Priorities/Sequencing	Instructor
	Learning Context Influences	Institutional Environment	Learner

Results from Semi-structured Interviews

After analyzing the data from CPs and PMTs, the ESP instructors and learners' interview transcripts were independently subjected to a similar thematic analytical procedure. Tables A3 and A4 in the Appendix detail the thematic analysis results based on the ESP instructors and learners' interview data, respectively. A synthesis of the results from interview data analysis is shown in Table 4.

Table 4. *Thematic Analysis Results based on the Interview Data*

Ctg.	Theme	Subtheme	Stakeholder
Pedagogical /Learning Needs	Dig. Experience & Tools Usage	Initial AI Adoption	Instructor
	Dig. Experience & Tools Usage	Informal Early AI Use	Learner
	Dig. Experience & Tools Usage	Familiarity with Digital Tools	Instructor
	AI Design Expectations	Preferred AI Features	Instructor
	AI Design Expectations	AI Adoption Requirements	Instructor
	AI Design Expectations	Field-Specific Functionality	Learner
	AI Design Expectations	Transparent, Justified Feedback	Learner
	AI Use Chl. & Opp.	Personalized Learning Potential	Learner
	Instructional/Ethical Roles	Ethical Tenets	Instructor
	Instructional/Ethical Roles	Ethical Transparency and Equity	Learner
	Instructional/Ethical Roles	Human-governed Tasks	Instructor
	Conceptual Understanding of AI	AI Application in ESP	Instructor
	Conceptual Understanding of AI	AI as a Learning Facilitator	Learner
AI Integration Perceptions	Conceptual Understanding of AI	AI Definition in Education	Instructor
	Conceptual Understanding of AI	AI Definition in Learning	Learner
	Instructional/Ethical Roles	Evolving Instructor/Learner Roles	Instructor
	Instructional/Ethical Roles	Instructors as Facilitators	Learner
	AI Use Chl. & Opp.	AI Supportive Role	Instructor
	AI Use Chl. & Opp.	Overdependence & Errors Risks	Learner
	AI Use Chl. & Opp.	Academic Integrity & Plagiarism	Learner
	AI Design Expectations	AI Acceptance Factors	Instructor
Contextual Challenges	AI Use Chl. & Opp.	AI Use Risks / Challenges	Instructor
	Dig. Experience & Tools Usage	Diverse Digital Literacy	Learner
	AI Design Expectations	Institutional Policy & Training	Learner
	Instructional/Ethical Roles	Human-led Assessment	Learner

Note. *Ctg. stands for "category", Dig. stands for digital, and Chl. & Opp stands for challenges and opportunities*

As shown in Table 4, under Needs, the instructors highlighted their familiarity with digital tools, early stages of AI adoption, and expectations for AI design, stressing preferred features, transparency, pedagogical suitability, and ethical responsibilities such as preserving human-led decision-making. The learners emphasized AI-specific interfaces, personalized learning abilities, and fair feedback, while both groups showed changing perspectives on AI, with the instructors stressing domain-specific usefulness and the learners perceiving

AI as enhancing their learning. Regarding AI integration perceptions, the instructors focused on AI’s evolving role from content delivery to facilitation and on factors shaping its acceptance, including relevance, ethical compatibility, and instructional suitability. The learners echoed this shift, viewing instructors as guides but also raising concerns over overreliance, error-prone outputs, and risks to academic integrity. Finally, under Contextual Challenges, the instructors discussed broader dangers such as unintended consequences and misuse, whereas the learners underscored digital literacy gaps, institutional constraints, training needs, and ethical issues in human-led assessment, calling for transparent and fair AI use in evaluation.

Synthesized Results

Figure 2 provides a thematic response to the first research question by integrating the results from different data resources (CPs, PMTs, and interviews). The shift from source-specific results in Tables 3 and 6 to the synthesized, streamlined thematic structure in Figure 2 entailed consolidating overlapping subthemes into broader categories while preserving distinct insights.

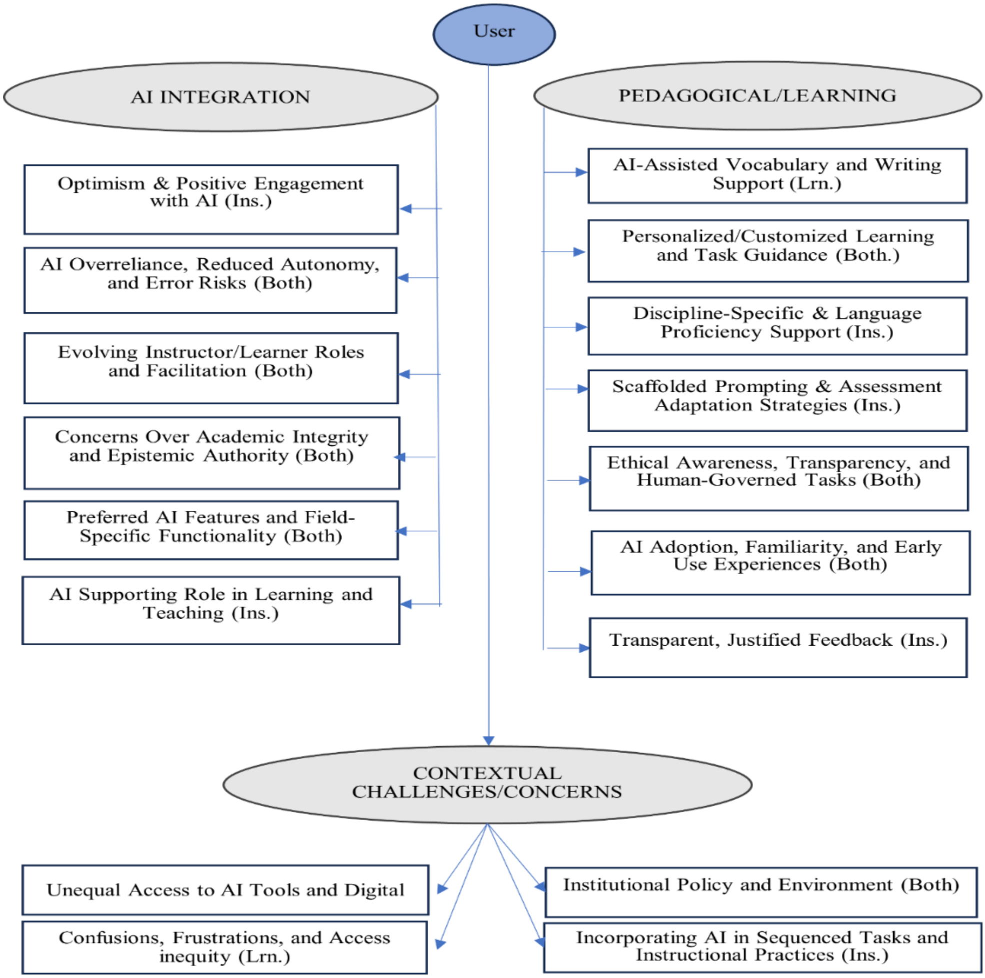


Figure 2. *Thematic Categorization of ESP Stakeholders' Data in Response to Research Question One*

Addressing the Second Research Question

Workshop 1 Results

The diverse data sets gathered in Workshop 1 fell into two main categories: textual data from RICs and ECCs, and visual data, including collaboratively created flow diagrams and storyboards using paper markers and handwritten annotations. The dual analytical framework combined inductive thematic analysis of textual data with artifact-based coding of visual materials, addressing both symbolic representations (markers, annotations) and functional instructional sequences (AI initiation, review gates, feedback loops). Additionally, observational field notes and memos on group dynamics and decision-making processes were used to enhance the interpretive validity and contextual richness of the emergent themes. The results are presented in the following sub-sections.

Role-specific Needs

Analyzing the data from RICs resulted in eight core themes (see Table A5 in the Appendix). The most prominent theme, Explicit Human-Check, was cited mainly by the ESP policymakers and the instructors, who underscored the need for a verification stage between AI output generation and subsequent instructional use to safeguard accuracy, relevance, and learner cognitive engagement. The next most frequent themes included Scaffolded Prompting Strategies and Assessment Rubrics Fostering Engagement. The former, voiced mostly by the ESP instructors and the learners, highlighted the need for teacher-mediated prompt design to ensure semantic precision and alignment with ESP objectives, whereas the latter mainly reflected the instructors' emphasis on cognitive processes beyond surface fluency to counter the risk of superficial output acceptance. Less frequent but noteworthy themes included Transparent AI Decision Paths, Context-Specific Resource Repositories, Continuous Teacher Training, Real-time Error Flagging, Adaptive Feedback Loops, and Integration of Multimodal Inputs.

Role-specific Concerns

As Table A6 in the Appendix shows, overreliance was the most noteworthy concern of learners, instructors, and policymakers, who strongly believed that heavy use of AI-produced content would inevitably reduce learner autonomy and critical reflection. Bias in AI output was another concern voiced mostly by AI experts and ESP instructors, who underscored challenges that arise when the tool produces culturally or contextually misaligned outputs. Another key concern of policymakers and learners was infrastructure challenges, such as inconsistent connectivity and unreliable hardware, which can hinder the equitable adoption of AI technologies. Privacy and confidentiality risks were the next concern area for ESP learners, who felt worried about unauthorized access to or misuse of sensitive data. Concerns voiced less frequently included language drift from ESP terminology, erosion of instructor authority, and ethical ambiguity in AI feedback.

Role-specific Suggestions

The thematic scheme that emerged from analyzing the RICs data (see Table A7 in the Appendix) resulted in several recommendations. The most noteworthy one made by the policymakers, instructors, and AI experts was to implement a formal human-check gate to ensure systematic verification of AI-generated content before learner submission, thereby maintaining instructional quality and ethical standards. The instructors and learners mostly emphasized designing scaffolded prompt templates tailored to discipline-specific needs to foster precise and pedagogically aligned AI interactions. The policymakers and instructors also recommended having proper regard for continuous teacher training programs aimed at enhancing educators' AI literacy and their capacity to evaluate AI outputs critically. Some of the AI experts and instructors suggested having foresight to enhance AI decision-making transparency, emphasizing that explainability is crucial for building user trust. Furthermore, the learners and instructors jointly proposed that context-specific resource repositories and adaptive feedback mechanisms need to be designed to support personalized learning pathways. As another suggestion, a couple of the policymakers and learners suggested developing robust data privacy protocols to protect sensitive learner information within AI-enabled educational settings.

Ethical Priorities

Collectively working in mixed-role groups and using ECCs, the participants identified nine prominent cultural concerns grouped under four core themes.

The groups proposed several mitigations to overcome these core concerns (see Table 5).

Table 5. *Ethical Priorities and Mitigation Strategies*

Theme	Code	Mitigation Strategy
Learner Data Privacy	Data Storage and Privacy	• Adopting LMS-integrated AI tools to prevent external data sharing • Anonymizing data before processing; securing explicit informed user consent.
	Unauthorized Data Access	• Implementing strict access controls and role-based management • Encrypting data during transmission and storage.
	External Server Risks	• Enforcing institutional security policies • Conducting regular security audits on external data handling.
Academic Integrity	Uncritical Use and Plagiarism	• Establishing human-check gates and enforce academic honesty policies • Training users on ethical AI use.
	Source Attribution and Critical Thinking	• Providing critical thinking and information literacy training • Clarifying plagiarism policies relating to AI-generated content.
Algorithm Bias	Bias and Fairness in AI Outputs	• Conducting regular audits and transparency mandates to detect and address AI bias • Educating users on recognizing and mitigating bias.
Other Concerns	Lack of Transparency in AI Functioning	• Developing user-friendly documentation and visual explanations of AI processes • Offering frequently asked questions and ongoing user education.
	Digital Divide and Access Inequality	• Implementing equitable access policies • providing offline or low-resource AI tool versions • Addressing digital literacy gaps.
	Security of Institutional Data	• Deploying advanced security protocols, continuous monitoring, and automated encryption for sensitive institutional data protection.

Scenario Storyboards

Thematic analysis of the ideal scenarios developed by the three groups revealed four core themes that reflect factors enhancing the effectiveness of AI integration. The problematic AI integration scenarios were also analyzed, and the resulting thematic scheme comprised four core themes that highlight challenges and risks hindering the effectiveness of the pedagogical model. Table 6 presents the results.

Table 6. *Thematic Scheme Drawn from Scenario Storyboard Analysis*

Scenario Type	Theme	Explanation
Ideal AI Integration	Structured Scaffolding via Prompt Design	Starting with clear and collaborative prompt setting ensures alignment with disciplinary expectations and supports learner autonomy
	Learner Agency and Engagement	Successful outcomes correlate strongly with active learner involvement in revising AI-generated drafts informed by expert feedback
	Crucial Role of Human Oversight	Implementation of human-check gates post-AI drafting is vital to intercept and amend errors, securing output validity and maintaining academic integrity
	Collaborative Iterative Revision Cycles	Dynamic triadic exchanges among teacher, learner, and AI promote metacognitive growth, critical thinking, and content mastery
Problematic AI integration	Neglected Human-Check Controls	Underuse or misapplication of review stages results in passing unverified AI content, propagating errors and dependency on AI
	Insufficient AI Accessibility and Competency	Limited AI tools availability or inadequate user training exacerbates learner confusion and suboptimal draft quality
	Potential Cognitive Decline	Excessive reliance on AI risks diminishing learners' independent analytical capacities and linguistic skills
	Ethical and Equity Concerns	Privacy protection, fairness in access to AI resources, and maintaining originality are pervasive concerns requiring institutional policy adaptation

Instructional Flowcharts

The three instructional flowcharts plotted by the three mixed-role groups over Workshop 1, each capturing the unique instructional perspectives and priorities of its respective participants, were analyzed carefully and synthesized based on a comparative thematic analysis to develop a prototype for AI-enhanced ESP instruction flow. To this end, every single flowchart was examined meticulously by the researcher to detect both commonalities and group-specific variations in sequences, actions, and decision points. The shared actions, roles, and phases constituted the backbone of the prototype instructional model, whereas the diverging ones, either those showing conceptual overlapping or those representing a unique role, action, or flow, were regarded as consensus-making and discussion topics over Workshop 2. The prototype instructional flow is displayed in Figure 3.

As illustrated in Figure 3, the instructional flow represents a cyclical process incorporating technological affordances into pedagogical routines. Informed by continuous professional developments and instructional supports, the iterative process commenced with prompt design and task set up, where ESP instructors and learners collaboratively embark on two successive activities: first, framing guiding prompts based on the targeted disciplinary context, linguistic focus, and task outcomes, and second, embedding the generated prompts into definite learning tasks tailored to the target ESP activities, such as developing discipline-specific problem-solving tasks, professional texts, and workplace communication scenarios. The process then continues with AI draft generation, where the AI generates the initial sketch of the required output. The draft is evaluated instantly, going through a human-check gate, where peers or ESP instructors assess the relevance and accuracy of the AI-generated output. Allowing learners to refine the text through iterative modification and resubmission, this evaluation and verification loop continues until the draft passes the human review stage. The process then advances as the verified draft moves to final submission by ESP learners. What concludes the process is the instructor assessment and feedback stage, in which learners receive feedback that not only assesses their performance but also guides the design of future prompts and shapes their engagement with tasks.

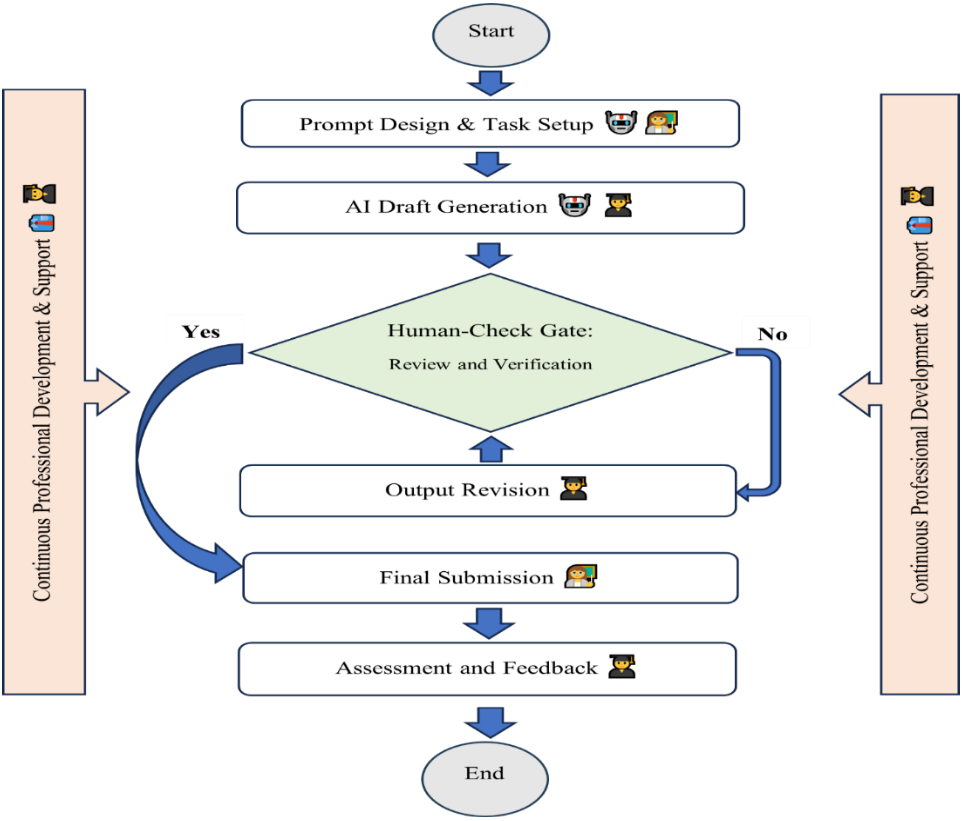
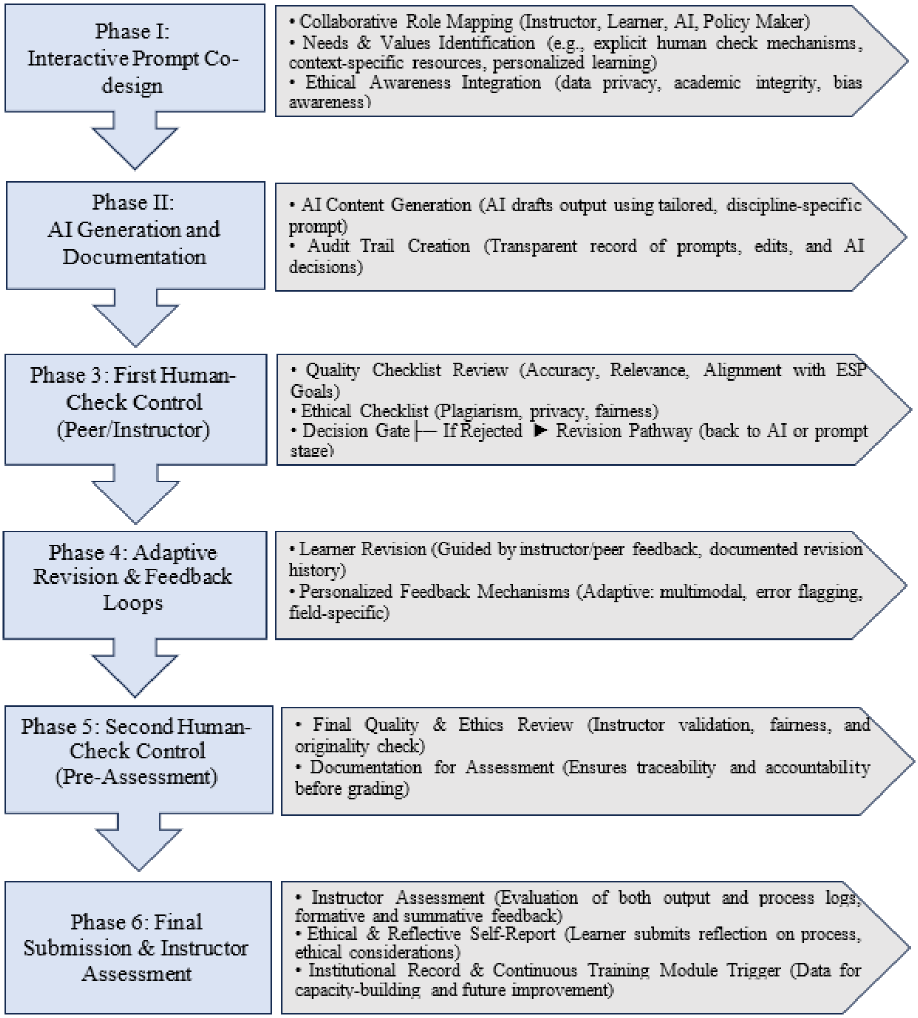


Figure 3. *Thematic Categorization of ESP Stakeholders' Data in Response to Research Question One*

Based on the prototype instructional flow and the thematic schemes that emerged from analyzing IRCs, TCCs, and SSTs, a prototype model for AI-assisted ESP instruction was developed (see Figure 4). Comprising six interrelated phases, the prototype model represented a cyclical pedagogical workflow combining the expertise of instructors, the agency of learners, and the assistive capabilities of generative AI.



Note. *The cycle repeats for iterative improvement and further tasks*

Figure 4. *Prototype Model for AI-assisted ESP Instruction*

Workshops 2 and 3 Results

The results drawn from the second and third workshops, which were intended to refine and finalize the prototype model, respectively, are displayed in Table 7.

Table 7. *Results Drawn from Workshop 2 Data Analysis*

	Strand	Theme	Refinement
Workshop 2	Expansion	Scaffolding Revision Cycles	Minimum two iterative cycles; checklists for content, language, and discipline-specific accuracy
		Ethical Literacy Integration	Micro-reflections integrated within revision loop; aligned with academic integrity rubric
		Audit Trails & Transparency	New Audit Trail (Explainability Record) node; dotted links to learner revision and instructor review
	Ambiguities	Criteria for Reflection	Clarification (no structural change) → Triggered at each major revision and at final submission; prompts standardized
		Teacher vs. AI Facilitation	Clarification (no structural change) → AI = generative support; teacher = interpretive oversight (explicit labels in model)
	Divergences	Scope of Human Oversight (single vs. dual human-check)	Agreement on dual human-check with evaluation-mapped criteria; time-cost mitigated via lightweight peer gate + targeted instructor gate
		Ethical Self-Reporting	Embedded micro-reflections inside the revision loop (≤100 words) instead of a standalone task
		Placement & Sequencing of Review Checkpoints	Final sequence set to: AI draft → Peer gate → Learner revision → Instructor gate → Final submission
Workshop 3	Institutionalization	Teacher Training & Record-Keeping	Added Institutional Record & Continuous Teacher Training node branching from Instructor Assessment; standardized reporting templates; alignment with departmental QA cycles
	Ethical Safeguards	Bias Awareness & Fairness	Inserted Bias Awareness & Fairness Audit gate immediately before Instructor Assessment; bias-note field added to Audit Trail
	Learner Empowerment	Agency via Iterative Revision	Retained and codified minimum two revision cycles; rubric dimension on independent reasoning and source triangulation
	Pedagogical Enhancement	Multimodal Feedback Loops	Expanded Assessment & Feedback to branch into Text / Audio / Visual feedback; learner-preferred modality recorded in audit trail; accessibility noted

The data drawn from the review, refinement, and expansion of the initial prototype in Workshop 2 were structured around three main strands: divergencies carried over the first workshop, potential areas for expansion, and ambiguities in task boundaries and role definitions. On the other hand, the finalized comments made to maximize the contextual relevance and practicality of the refined model in Workshop 3 were grouped into four groups, including institutionalization, ethical safeguards, learner empowerment, and pedagogical enhancement. Table 7 provides a summary of the thematic results drawn from these workshops and the changes made in the initial and refined model.

The improvements from Workshop 2, specifically activating both human-check gates in sequence, weaving micro-revisions into existing cycles, and establishing clear audit trails, were folded into the polished prototype model. This model served as the basis for Workshop 3, convened as a consensus forum to validate and finalize the design. Results from this concluding phase, distilled into four thematic priorities, are reported in Table 7. The consensus session affirmed institutional viability and identified four subsequent mandates: institutional record-keeping and teacher development; a formal checkpoint for bias-awareness; the safeguard of learner agency via at least two revision cycles; and the delivery of multimodal feedback. The approved directives were operationalized in the production model by embedding an institutional node (records/training), anonymizing learner data before instructor assessment, maintaining the iterative revision cycle with rubric-based agency criteria, and branching output into text, audio, and visual forms. These cumulative modifications converged into the finalized production model, as illustrated in Figure 5

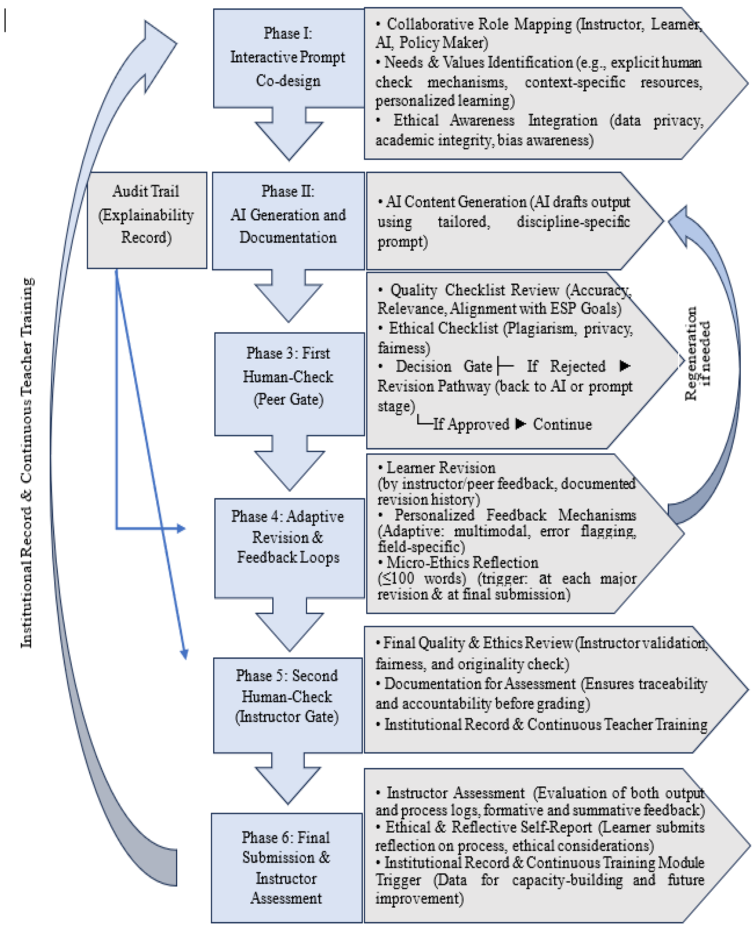


Figure 5. *Finalized Model for AI-assisted ESP Instruction (Post-Workshop 3)*

DISCUSSION

The present study pursued two interrelated aims: (a) to map Iranian ESP instructors' and learners' perceptions, needs, and contextual challenges regarding generative AI integration, and (b) to co-design a contextualized, sustainable pedagogical framework through intensive stakeholder participation. The discussion is therefore structured around these two objectives.

Perceptions, Needs, and Contextual Challenges in AI-enhanced Iranian ESP Education

The analysis surfaced two distinct but mutually illuminating sets of views among the ESP students and their instructors. Learners oriented toward the application of generative AI sought individualized micro-scaffolding contextualized to specific disciplines, along with explicit writing support, regarding the technology primarily as a tactical resource for immediate gains. This behavior substantiates claims that AI-centric micro-adaptive scaffolding is a core pedagogical asset (Nguyen, 2025; Weng & Fu, 2025) and harmonizes with sociocultural emphases on the ZPD concept in situated contexts (Lantolf & Thorne, 2006; Vygotsky, 1978). Furthermore, it is in line with nascent studies that argue discipline-specific tailoring can amplify learner autonomy (Elmotri et al., 2025). In contrast, the ESP instructors foregrounded controlled prompting, principled assessment reconceptualization, and a developing ethics literacy agenda as mechanisms to preserve academic integrity and pedagogical alignment. Their caution tracks well-documented threats of automation bias and epistemic unreliability, as well as the potential for both students and educators to lose agency (Artyukhov et al., 2025; Luckin et al., 2016). The disparity between the ESP instructors and learners corroborated the empirically validated (e.g., Petricini et al., 2025; Zipf et al., 2025) role-specific perceptions in ESP settings, showing that, unlike instructors' high regard for institutional and epistemic gains, learners mostly prioritized their immediate learning attainments.

On a wider canvas, the findings revealed the distinctly situated trajectory of AI in Iranian ESP instruction. As evidenced by prior evidential data (e.g., Hosseini Moghadam, 2023; Mostafavi et al., 2021), the learners' pleas for learning personalization and customization accentuated persistent aspirational intimidation for flexibility within entrenched, exam-centric epochs, whereas the instructors' demands for ethical literacy signal a hesitance to repose confidence in structures that are already plausibly fissured by concerns over authorship and credential integrity. Likewise, the issue of differential access echoed global debates on the digital divide (Dolba et al., 2023; Pawar & Khose, 2024). However, the concern may take on an even sharper resonance in the Iranian context of infrastructural fragility (Hosseini Moghadam, 2023; Mostafavi et al., 2021).

Notwithstanding the differences in role-specific needs and challenges, the views of stakeholder groups converged on the potential risks of epistemic authority negotiation, AI overreliance, and the persistence of digital inequity. These shared concerns, however, were voiced from varying viewpoints by the two stakeholder groups. The learners were concerned about the detrimental role of AI dependence in the knowledge mastery path, whereas the instructors' concerns were rooted in their concern for their eventual loss of authority. This asymmetrical concern, despite shared ethical considerations, necessitates a multifaceted AI integration process. Such an approach counters the hazard densely emphasized in the literature: technology adoption without due regard for risks and role-specific nuances could exacerbate fragmentation and reinforce inequality (Luckin et al., 2016).

Overall, the context-specific needs, perceptions, and challenges articulated by key stakeholders of an AI-enhanced ESP pedagogy in the Iranian academic setting were largely influenced by context-driven systematic inequities and professional vulnerabilities. It seems quite reasonable to acknowledge that AI adoption demands and challenges in the specific context of Iranian ESP pedagogy are hardly compatible with the findings of previous studies in settings (e.g., Nguyen, 2025; Wu, 2024), where digital access is widespread and infrastructures can easily meet the ongoing demands of such a blended instructional approach. The literature is replete with comparative evidence showing that while frameworks in Europe and East Asia prioritize innovation, this region foregrounds ethical governance and equity, offering transferable lessons for AI integration in the Global South (Giannakos et al., 2024; Holstein et al., 2019). All in all, aligned with evidential data (e.g., Marchena Sekli et al., 2024; Zaim et al., 2025) that emphasize due regard for contextual peculiarities and stakeholders' reflective engagement while conceptualizing and implementing AI-assisted education as a safeguard against superficial adoption, the study's findings emphasize on co-designing a context-relevant model of AI-assisted ESP pedagogy.

Co-Designed Model for AI-Assisted ESP Instruction

As its second inquiry domain, the study explored an AI-enhanced ESP instructional model contextually suited to the Iranian higher education context. The emergent model proved to be anchored on recursive phases where all components influenced and adjusted to one another in an iterative loop. The first phase involved the joint crafting of task prompts, where instructors and learners are to articulate performance criteria together and select disciplinary content and language features while framing the AI as a partner lifted by a broad co-leadership in task design. Next, the system generates an initial version of content, obliging the AI to log its reasoning, thus creating a map of inference decisions for future scrutiny, a safeguard known as the explainability portfolio. Two human checkpoints follow in sequence: a peer revision round that encourages collaborative diagnosis, redistributing cognitive effort and tempering reliance on the machine, and a teacher approval phase that guarantees the content's disciplinary soundness and its fitness to institutional ethics criteria. Between these checkpoints, iterative drafts enter an adaptive revision phase, in which students polish texts with targeted multimodal feedback, such as graphics, sound annotations, or video micro-lessons, selected by each learner's operational preference. Embedded in each version is a micro-ethical question, nudging the writer to confront issues of originality, algorithmic bias, and equitable representation of represented groups. The loop is finally concluded by institutional archiving of outputs and by real-time updates to the training database for teachers, thus ensuring that the design can embed, endure, and evolve within the university's evolving policy.

An accurate reflection on this pedagogical route could imply its complex nature, encompassing both theoretical concepts and embodied experiences. The first stage, co-constructed prompt design, proved consistent with the SCT, acknowledging the scholarly contention that adopting GenAI integrates AI-mediated learning into learners' ZPD through collaborative inquiry and guided mediation to enhance cognitive and affective engagement (Roshan et al., 2024; Shrestha et al., 2025). On the other hand, the dual human-check cycle solidifies the tenets of TPACK, integrating state-of-the-art technological affordances, subtle paths of sound pedagogy, and discipline-specific knowledge hints. The triangulation of knowledge resources, as noted by Mishra and Koehler (2006), can help ensure that the output of any working model demonstrates both fluency and accuracy.

One probable justification for the model's practicality in the Iranian ESP context may lie in achieving a working compromise between the systemic restrictions faced and the educational objectives set. Due to infrastructural deficiencies, centralized exams, and differences in digital literacy, the use of AI is not thoroughly viable as an educational tool within Iran's education system. AI's limitations, combined with the dual human-check cycle, audit trails, and ethical self-reporting, could alleviate some of the disadvantages associated with inequality while maintaining reliability. The model also offered other potential benefits to consider. For instance, through formal teacher training and multi-modal feedback systems, gaps in learner accessibility and professional preparedness in the Iranian ESP context (Ghiasvand et al., 2024; Roshan et al., 2024; Shrestha et al., 2025) can be bridged, resulting in systemic resiliency and long-term sustainability.

Adaptive revision, carried out in multimodal circuits of feedback, preserves space for learners' own metacognitive monitoring and fills the gaps in contemporary digital literacies often identified in the existing literature (e.g., Ghiasvand et al., 2024; Holstein et al., 2019). Finally, ethical micro-reflections and audit trails embedded in the model iteratively operationalize human-centered AI principles, fostering learners' critical interrogation of generative horizons and the development of ethical fluency within clear landscapes (Derakhshan et al., 2025; Pawar & Khose, 2024; Zhyhadlo & Zaiarna, 2025; Zipf et al., 2025). Institutionalization through continuous teacher training and record-keeping may also embed the model into organizational practice, aligning it with previous literature that supports the efficacy of such practices (e.g., Dolba et al., 2023; Luckin et al., 2016; Maghferat et al., 2024; Parviz, 2024) and alters it from a classroom-level experiment into a systemic framework adaptable to similar ESP contexts worldwide.

The co-designed, multi-layered model, which emerged to shed light on AI integration in an Iranian ESP context, may offer critical pedagogical, theoretical, and institutional implications. Pedagogically, within this model, learners are positioned as co-agents. Nevertheless, instructor authority is preserved, as the integrated AI-supported system employs dual human oversight, adaptive revisions, and multimodal feedback to enhance domain-specific competencies, foster learner autonomy, and mitigate plagiarism or epistemic errors, emphasizing process-oriented assessment aligned with ESP goals (Holstein et al., 2019; Weng et al., 2024; Xia et al., 2024). Theoretically, it mobilizes the SCT and TPACK through a combination of scaffolding, iterative "prompting," micro-reflections, and audit trails, whereby a dual-layer feedback and gatekeeping model, ethically framed and embedded in both the learning and institutional fabric, provides feedback within learning contexts (Mishra & Koehler, 2006; Zhyhadlo & Zaiarna, 2025). The model also appears compatible with the Iranian institutional context, as it accounted for AI literacy and ethical training within professional teaching frameworks, executing equity-focused infrastructure, and addressing digital divides in Iranian ESP contexts, as the context-specific peculiarities of the local context under investigation (Ghiasvand et al., 2024; Roshan et al., 2024; Shrestha et al., 2025).

CONCLUSION AND IMPLICATIONS

This research developed a context-specific, co-designed framework for the application of GenAI in ESP instruction in Iran, ensuring the approach is pedagogically sound, ethically sensitive, and institutionally sustainable. The model constructed through stakeholder participation included prompt scaffolding, dual human-check gates, ethical self-reflective revision cycles, and multimodal feedback coupled with ongoing teacher training, thus integrating innovation with human and ethical dimensions. By emphasizing learner autonomy, instructor guidance, and institutional infrastructure, the research illustrated that the effective integration of GenAI in ESP is not a function of technological features alone, but requires an intentional blending of pedagogy, ethics, and context. The research offered a comprehensive practical blueprint for Iranian ESP teaching while providing a model for similar educational settings attempting to integrate AI technologies in a context-sensitive manner.

While proceeding to the final model, the research faced several limitations. The focused sample drawn from four branches of Islamic Azad University overlooked broader perspectives. The qualitative, design-based approach to data gathering limited a comprehensive assessment of learning outcomes. The use of self-reported data and participatory artifacts may have resulted in biased interpretations. Further studies may focus on longitudinal studies on classroom trials of the model, studies on incorporating AI into developing assessment methods and equity-focused initiatives, cross-institutional studies to test the model's scalability, and comparative studies contrasting the model's components and flow with those suited to the Global North contexts. These scientific endeavors would provide authentic evidence of the model's effectiveness, help adapt it for different contexts, and support the delivery of ESP education in an inclusive and ethically responsible manner through the integration of AI.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Masoud Azadnia http://orcid.org/0000-0001-5425-7862

References

AbuSahyon, A. S., Alzyoud, A., Alshorman, O., & Al-Absi, B. (2023). AI-driven technology and chatbots as tools for enhancing English language learning in the context of second language acquisition: A review study. *International Journal of Membrane Science and Technology, 10*(1), 1209–1223. https://doi.org/10.15379/ijmst.v10i1.2829

Adam, I. O. (2020). Using a qualitative interpretive approach in educational technology implementation: A personal experience from a developing country university. In E. Boateng & K. Effah (Eds.), *Handbook of research on managing information systems in developing economies* (pp. 318–321). IGI Global.
https://doi.org/10.4018/978-1-7998-2610-1.ch015

Ahmed, Z., Qazi, U., & Raza, M. (2025). Exploring the role of artificial intelligence for revolutionizing English language learning of university-level students. *ACADEMIA: International Journal for Social Sciences, 4*(3), 27–38. https://doi.org/10.63056/ACAD.004.03.0333

Artyukhov, A., Artyukhova, N., Dluhopolskyi, O., Adamyk, O., & Adamyk, B. (2025). Dialogue with generative artificial intelligence: Is its "product" free from academic integrity violations? Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, (2), 181–188. https://doi.org/10.33271/nvngu/2025-2/181

Atai, M. R., & Tahririan, M. H. (2003). *Assessment of the status of ESP in the current Iranian higher educational system*. In Proceedings of the 14th European Symposium on Language for Special Purposes: Communication, culture, and knowledge. University of Surrey, Guildford, United Kingdom.

Bagnoli, A. (2009). On "an introspective journey": Identities and visual artefacts in youths' transition to adulthood. *International Journal of Social Research Methodology, 12*(2), 167–180. https://doi.org/10.1080/13645570802616540

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Dashtestani, R., & Stojković, N. (2016). The use of technology in English for specific purposes (ESP) instruction: A literature review. *Journal of Teaching English for Specific and Academic Purposes, 3*, 435–456.

Demirdöken, G. (2024). Artificial intelligence in aviation English testing. *The Literacy Trek*, *10*(3), 362–384. https://doi.org/10.47216/literacytrek.1556603

Derakhshan, A., Karimpour, S., & Nazari, M. (2025). Exploring the professional role identities of English for academic purposes practitioners: a qualitative study. *International Review of Applied Linguistics in Language Teaching*, 63(2), 1295–1316. https://doi.org/10.1515/iral-2023-0126

Dolba, S. Q., Inoncillo, F. A., & Nunez, J. L. (2023). Evaluating the impact of AI-generated outputs on student assessment: Educator's perspective. *International Journal of AI Research*, *9*(1), 1–7. https://doi.org/10.29099/ijair.v9i1.1347

Elmotri, B., Osman, W. A. E. A. E., Al-humari, M. A., Amri, F., Elyasa, Y. M., Malik, F. H., Garrouri, S., & Hassanein, O. H. S. (2025). Possibilities for improving ESP curriculum design and assessment strategies for Saudi universities with ChatGPT. *Arab World English Journal (AWEJ), Special Issue on Artificial Intelligence*, 2025, 234–250. https://dx.doi.org/10.24093/awej/AI.13

Eslami, Z. (2010). Teachers' voice vs. students' voice: A needs analysis approach of English for academic purposes (EAP) in Iran. *English Language Teaching, 3*(1), 3–11. https://doi.org/10.5539/elt.v3n1p3

Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science, 81,1–9.* https://doi.org/10.3389/bjbs.2024.14048

Ghiasvand, F., Kogani, M., & Alipoor, A. (2024). I'm not ready for this metamorphosis: An ecological approach to Iranian and Italian EFL teachers' readiness for artificial intelligence-mediated instruction. *Teaching English with Technology*, *24*(3), 18–40. https://doi.org/10.56297/vaca6841//BFFO7057/ISKI2001

Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative AI in education. *Behaviour & Information Technology*, *44*(11), 2518–2544. https://doi.org/10.1080/0144929X.2024.2394886

Holmlid, S. (2009). Participative, co-operative, emancipatory: From participatory design to service design. *CoDesign*, *5*(2), 79–83. https://doi.org/10.1080/15710880903012769

Holstein, K., McLaren, B. M., & Aleven, V. (2019). Designing for complementarity: Teacher and student needs for orchestration support in AI-enhanced classrooms. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), Artificial Intelligence in Education (pp. 157–171). Springer. https://doi.org/10.1007/978-3-030-23204-7_14

Hosseini Moghadam, M. (2023). Artificial intelligence and the future of university education in Iran. *Research and Planning in Higher Education Quarterly, 29*(1), 1–25. https://doi.org/10.61838/irphe.29.1.1

Jegede, O. O. (2024). Artificial intelligence and English language learning: Exploring the roles of AI-driven tools in personalizing learning and providing instant feedback. *Universal Library of Languages and Literatures*, *1*(2), 6–19. https://doi.org/10.70315/uloap.ullli.2024.0102002

Jia, C. (2025). The impact of curriculum design and educational equity in the age of artificial intelligence: A literature review. In *Proceedings of the 6th International Conference on Education Innovation and Psychological Insights* (pp. 55–61). https://doi.org/10.54254/2753-7048/2025.22633

Krishnan, V., & Zaini, H. (2025). A systematic literature review on artificial intelligence in English language education. *International Journal of Research and Innovation in Social Science (IJRISS), IX*(IIIS), 17–27. https://dx.doi.org/10.47772/IJRISS.2025.903SEDU0002

Lai, Z. C.-C. (2025). The impact of AI-assisted blended learning on writing efficacy and resilience. *International Journal of Computer-Assisted Language Learning and Teaching*, *15*(1), 1 –21. https://doi.org/10.4018/IJCALLT.377174

Lantolf, J. P., & Thorne, S. L. (2006). *Sociocultural theory and the genesis of second language development*. Oxford University Press.

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson. https://doi.org/10.17863/CAM.39544

Maghsoudi, M. (2023). Factors determining the technological pedagogical content knowledge (TPACK) of the Iranian preservice English teachers. *Issues in Language Teaching*, *12*(2), 1–30. https://doi.org/10.22054/ilt.2023.72069.754

Marchena Sekli, G. F., Godo, A., & Véliz, J. C. (2024). Generative AI solutions for faculty and students: A review of literature and roadmap for future research. *Journal of Information Technology Education: Research*, 23, 1–23. https://doi.org/10.28945/5304

Mckenney, S., & Reeves, T.C. (2012). *Conducting educational design research*. Routledge.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Mostafaei Alaei, M., & Ershadi, A. R. (2016). ESP program in Iran: A stakeholder-based evaluation of the program's goal, methodology, and textbook. *Issues in Language Teaching*, *5*(2), 306–279. https://doi.org/10.22054/ilt.2017.8062

Mostafavi, S., Mohseni, A., & Abbasian, GR. (2021). The pedagogical efficacy of ESP courses for Iranian students of engineering from students' and instructors' perspectives. *Asian Pacific Journal of Second and Foreign Language Education. 6*(6), 1–20. https://doi.org/10.1186/s40862-021-00109-2

Nezakatgoo, B., & Behzadpoor, F. (2017). Challenges in teaching ESP at medical universities of Iran from ESP stakeholders' perspectives. *Iranian Journal of Applied Language Studies, 9*(2), 59–82. https://doi.org/10.22111/ijals.2017.3544

Nguyen, T. V. (2025). The impact of AI tools on English writing skills development. *International Journal of Social Science Exceptional Research*, *4*(3), 1–6. https://doi.org/10.54660/IJSSER.2025.4.3.01-06

Noroozi, O., Soleimani, S., Farrokhnia, M., & Banihashem, S. K. (2024). Generative AI in education: Pedagogical, theoretical, and methodological perspectives. *International Journal of Technology in Education (IJTE)*, *7*(3), 373–385. https://doi.org/10.46328/ijte.845

Pan, J. (2024). AI-driven English language learning program and academic writing integrity in the era of intelligent interface. *English Language Teaching and Linguistics Studies*, *6*(4), 120 –130. https://doi.org/10.22158/eltls.v6n4p120

Parviz M. (2024). The double-edged sword: AI integration in English language education from the perspectives of Iranian EFL instructors. *Complutense Journal of English Studies*, *32*, 1–19. https://doi.org/10.5209/cjes.97261

Pawar, G., & Khose, J. (2024). Exploring the role of artificial intelligence in enhancing equity and inclusion in education. *International Journal of Innovative Science and Research Technology, 9*(4), 2180–2185. https://doi.org/10.38124/ijisrt/IJISRT24APR1939

Petricini, T., Zipf, S., & Wu, C. (2025). RESEARCH-AI: Communicating academic honesty: Teacher messages and student perceptions about generative AI. *Frontiers in Communication*, *10*, 1–6. https://doi.org/10.3389/fcomm.2025.1544430

Puspasari, A., & Agustina, N. (2025). The integration of AI in English for nursing instruction: A systematic literature review. *Automatic Zoom: Journal of English Teaching, 11*(1), 50–65. https://doi.org/10.33541/jet.v11i1.6509

Rajabi, P., Kiany, G.R., & Maftoon, P. (2012). ESP in-service teacher training programs: Do they change Iranian teachers' beliefs, classroom practices and students' achievements? *Iberica*, 261–282.

Renfeng, J., Gang, Y., & Qi, S. (2025). The Motivational Impact of GenAI Tools in LanguageLearning: A Quasi-Experiment Study. *International Journal of Applied Linguistics*, *35*(3), 1338-1350. https://doi.org/10.1111/ijal.12701

Rose, G. (2016). *Visual methodologies: An introduction to researching with visual materials* (4th ed.). Sage Publications.

Roshan, S., Iqbal, S. Z., & Qing, Z. (2024). Teacher training and professional development for implementing AI-based educational tools. *Journal of Asian Development Studies*, *13*(2), 1972–1987. https://doi.org/10.62345/jads.2024.13.2.154

Sanders, E. B.-N., & Stappers, P. J. (2014). Probes, toolkits and prototypes: Three approaches to making in codesigning. *Co-Design*, *10*(1), 5–14. https://doi.org/10.1080/15710882.2014.888183

Shakibaei, G., Soozandehfar, S. M. A., Owliaei, F., & Hashemifardnia, A. (in press). Unboxing practical strategies in university curriculum in the light of artificial intelligence: An insight into the effects on intellectual awareness, emotional intelligence, motivation, and creativity. *Issues in Language Teaching*. https://doi.org/10.22054/ilt.2024.80332.859

Shrestha, B. L. L., Dahal, N., Hasan, M. K., Paudel, S., & Kapar, H. (2025). Generative AI on professional development: A narrative inquiry using TPACK framework. Frontiers in Education, *10*, 1–9. https://doi.org/10.3389/feduc.2025.1550773

Smith, D. P., Sokoya, D., Moore, S., Okonkwo. C., Boyd, C., Lacey, M. M., & Francis, N. J. (2025). Embedding Generative AI as a digital capability into a year-long skills program. *Journal of University Teaching and Learning Practice*, *22* (3), 1–24. https://doi.org/10.53761/fh6q4v89

Spinuzzi, C. (2005). The methodology of participatory design. *Technical Communication*, *52*(2), 163–174.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Weng, X., XIA, Q., Gu, M., Rajaram, K., & Chiu, T. K. (2024). Assessment and learning outcomes for generative AI in higher education: A scoping review on current research status and trends. *Australasian Journal of Educational Technology*, *40*(6), 37–55. https://doi.org/10.14742/ajet.9540

Weng, Z., & Fu, Y. (2025). Generative AI in language education: Bridging divide and fostering inclusivity. *International Journal of Technology in Education (IJTE)*, *8*(2), 395–420. https://doi.org/10.46328/ijte.1056

Wu, Y. (2024). Study on the impact of utilizing ChatGPT and other AI tools for feedback in EAP writing classrooms on the discursive writing performance of English major students.*Transactions on Social Science, Education and Humanities Research*, *4*, 143–150. https://doi.org/10.62051/4se95x52

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, *21*(40), 1–22. https://doi.org/10.1186/s41239-024-00468-z

Zaim, M., Arsyad, S., Waluyo, B., Ardi, H., Al Hafizh, M., Zakiyah, M., Syafitri, W., Nusi, A., & Hardiah, M. (2025). Generative AI as a cognitive co-pilot in English language learning in higher education. *Education Sciences*, *15*(6), 686–411. https://doi.org/10.3390/educsci15060686

Zawacki-Richter, O., Marín, V.I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators?. *International Journal of Educational Technology in Higher Education*, *16*(39), 1–27.https://doi.org/10.1186/s41239-019-0171-0

Zhyhadlo, O., & Zaiarna, I. (2025). Artificial intelligence-driven testing in EFL/ESP classrooms. Information Technologies and Learning Tools, 106(2), 122-133. https://doi.org/10.33407/itlt.v106i2.5957

Zipf, S., Wu, C., & Petricini, T. (2025). Using the information inequity framework to study GenAI equity: analysis of educational perspectives. *Information Research an InternationalElectronic Journal*, *30*(iConf), 533–547. https://doi.org/10.47989/ir30iConf47284

Appendix

Table A1
Themes Emerged from ESP Instructors' Data on CPs and PMTs

Theme	Subtheme	Brief Explanation	Sample Extract
Pedagogical Routines	AI Tool Used in Routine Instructions	Embedding AI and educational technologies into classroom practices	Ins. 5: *"Students were using DeepL before I even mentioned translation; it's embedded now."*
	Task Priorities/Sequencies	Organizing ESP lessons and allocating time to skills/tasks	Ins. 2: *"I usually begin with a vocabulary warm-up, then use a sample text to discuss terminology"*.
Current AI Integration Practices	Vocabulary Development Practices	Using AI tools to support comprehension of technical vocabulary and definitions	Ins. 3: *"I used ChatGPT to help my students compare their definitions with AI-generated ones... they enjoyed critiquing the AI's word choices."*
	Prompted Writing Practices	Leveraging AI to initiate or rephrase ESP writing tasks	Ins. 8: *"Students prompted GPT to generate argument outlines... helped weaker students structure their thoughts."*
AI Use Challenges	Overreliance/Reduced Autonomy	Observing reduced student effort due to increasing dependence on AI	Ins. 2: *"Students relied on GrammarlyGO too heavily and didn't reflect on why their phrasing was off."*
	Unequal Access and Digital Literacy	Experiencing unequal access to AI-related devices or digital literacy	Ins. 1: *"A few students reported technical difficulties with app installation."*
	Doubt in Epistemic Authority	Questioning AI's reliability in domain-specific knowledge applications	Ins. 7: *"GPT wrote 'CRISPR cures bacteria,' which is incorrect... that was a wake-up call."*
Pedagogical Adaptations	Controlled Prompting Strategies	Designing explicit prompts to scaffold safe and pedagogically sound AI use	Ins. 6: *"I walked them through refining prompts and discussed how an 'architect's eye' is needed to fill in what AI misses."*
	Ethical Awareness Development	Gaining knowledge about issues like plagiarism, authorship, and bias	Ins. 4: *"We turned the task into a reflection: 'What did the AI miss that a real patient might need?' This opened up ethical dimensions."*
	Reframing Assessment Tasks	Shifting assessment focus toward learners' cognitive engagement over AI-generated fluency	Ins. 2: *"I revised the rubric on the spot to include 'authenticity of voice' after noticing over-dependence on rewriters."*

Table A2

Themes Emerged from ESP Learners' Data on CPs and PMTs

Theme	Subtheme	Brief Explanation	Sample Extract
Learning Supports	Writing Assistance	Employing AI to support learners in planning and refining written content	Lrn. 7: *"When I'm stuck, I ask AI to create sentence starters, which helps me write my ESP essays."*
	AI-based Vocabulary Aids	Leveraging AI for mastering technical terminology	Lrn. 3: *"I use Google Translate and ChatGPT to find exact meanings of difficult words in my microbiology texts."*
	Personalized/ Customized Learning	Facilitating need-based learning with AI-driven content personalization	Lrn. 12: *"I like how some apps adjust questions based on my mistakes, making learning more focused."*
AI Interaction Experiences	Confusions/ Frustrations	Facing challenges from AI misinterpretations, technical issues, or complex interfaces	Lrn. 5: *"Sometimes AI gives wrong definitions or mixes terms, which confuses me a lot."*
	Positive Engagement	Expressing satisfaction and enthusiasm toward AI-supported learning	Lrn. 9: *"Using GrammarlyGO made my assignments clearer, and I feel more confident submitting them."*
	Access Inequity	Facing limitations in AI access caused by varying connectivity and device availability	Lrn. 15: *"In my hometown, internet is slow, so I can't use some AI apps properly."*
Learning Context Influences	Discipline-Specific Needs	Adapting AI tools to meet the unique needs of diverse academic fields	Lrn. 6: *"In engineering, I need AI for formula explanations, but in psychology, I want help with essay structure."*
	Language Proficiency Support	Experiencing AI differently due to varying levels of English proficiency	Lrn. 14: *"I'm not confident in English, so I use AI more for translations and grammar help."*
	Institutional Environment	Guiding AI tool implementation through university-level decisions and infrastructure	Lrn. 1: "Our university doesn't allow some apps, so we have to find alternatives." (Participant 1)
AI Use Challenges	Overreliance on AI	Undermining critical thinking through overreliance on AI support	Lrn. 10: *"I noticed I often just copy what AI suggests without trying to understand it myself."*
	Accuracy Concerns	Showing uncertainty about the correctness of AI responses in specialized areas	Lrn. 18: *"AI sometimes gives wrong info about medical terms, so I double-check everything."*
	Digital Literacy Gaps	Reducing AI effectiveness through inadequate technical know-how	Lrn. 2: *"I don't always know how to phrase questions for AI to get good answers."*
Expectations/ Attitudes	Desire for Human-AI Balance	Favoring teacher-guided AI integration over full automation in education	Lrn. 13: *"AI is helpful, but I still want my teacher to explain things, especially tricky concepts."*
	Ethical and Academic Concerns	Highlighting learner concerns over fairness, misuse, and academic integrity	Lrn. 11: *"Sometimes I feel using AI might be cheating if I rely too much on it."*
	Optimism About AI Potential	Believing AI can greatly enhance ESP learning when effectively integrated	Lrn. 4: *"I think AI will change how we learn ESP, making it more interactive and interesting."*

Table A3

Themes Emerged from ESP Instructors' Interview Data

Theme	Subtheme	Brief Explanation	Sample Extract
Digital Experience & Tools Usage	Initial AI Adoption	Early-stage use of AI tools in instructional or learning activities	Ins. 8: *"I've encouraged AI-assisted revision tools like QuillBot and Claude for rewriting code documentation."*
	Familiarity with Digital Tools	Extent of use and comfort with digital tools such as LMS platforms, AI technologies, and multimedia resources	Ins. 1: *"I've used PowerPoint, Moodle, Grammarly, and occasionally Google Translate, but no full AI integration yet."*
Instructional/ Ethical Roles	Evolving Instructor/Learner Roles	Shift towards instructors as facilitators and learners as more independent.	Ins. 5: *"Teachers become facilitators of inquiry rather than content deliverers, which needs mindset shifts."*
	Human-governed Tasks	Tasks that require direct human oversight, such as assessments and ethical decision-making.	Ins. 4: *"Critical reasoning, peer evaluation, and ethical decision-making must remain human-centered."*
	Ethical Tenets	Emphasizing transparency, fairness, data privacy, academic integrity, and proper authorship	Ins. 1: "*Transparency, fairness, and data privacy are crucial; students should know AI's limitations and ethical use."*
AI Use Challenges / Opportunities	AI Supportive Role	How AI can enhance ESP learning through tools like vocabulary scaffolding and immediate feedback.	Ins. 7: *"AI boosts ideation and vocabulary fluency, especially in creative fields like architecture."*
	AI use Risks/Challenges	Issues such as over-reliance, plagiarism, misinformation, and ethical concerns.	Ins. 6: *"Bias in medical data and overgeneralization can mislead students."*
Conceptual Understanding of AI	AI Definition in Education	The interpretation of AI within the context of teaching and learning ESP	Ins. 3: *"AI means intelligent systems that personalize instruction or automate feedback like grammar checking."*
	AI Application in ESP	Instances of AI tool usage by teachers or students in ESP settings.	Ins. 2: *"Students use ChatGPT to write essays or generate ideas; I used it to prepare comprehension materials."*
AI Design Expectations	AI Acceptance Factors	Factors influencing trust in AI, including goal alignment, ease of use, transparency, and adequate training.	Ins. 8: *"I'd trust AI that allows feedback layering and integrates discipline-specific writing structures."*
	Preferred AI Features	Specific functionalities sought in AI tools designed for ESP disciplines.	Ins. 3: *"Interactive clinical case walkthroughs, smart glossary for Latin terms, plagiarism checker for medical writing."*
	AI Adoption Requirements	Critical design considerations addressing data privacy, authorship rights, misuse prevention, and safeguarding the teacher's role.	Ins. 2: *"I'd need clarity on data privacy, authorship, and teacher's role before full adoption."*

Table A4

Themes Emerged from ESP Learners' Interview Data

Theme	Subtheme	Brief Explanation	Sample Quote
Digital Experience & Tools Usage	Informal Early AI Use	Learners independently engaged with AI tools (e.g., ChatGPT, Grammarly) out of personal curiosity rather than formal instruction.	Lrn. 7: *"I first used ChatGPT to rewrite a lab report introduction. It was more about testing its limits than for an assignment."*
	Diverse Digital Literacy	Learners' varying degrees of experience and comfort with technology (AI tools	Lrn. 3: *"Before these classes, I mostly used online dictionaries and YouTube videos. AI tools seemed a bit too advanced for me."*
Instructional/ Ethical Roles	Instructors as Facilitators	Learners expect instructors to be mentors in digital literacy instead of traditional instructors	Lrn. 4: *"Our teacher showed us how to prompt better. That helped me trust the process more."*
	Human-led Assessment	Learners view oral exams and argumentative writing as requiring human evaluation.	Lrn. 8: *"AI can't really tell if an argument is strong or creative. Only teachers can judge that."*
	Ethical Transparency and Equity	Learners emphasize data privacy, disclosure of AI use, and addressing technology-driven inequalities	Lrn. 2: *"We should know if AI was involved in making materials—and if it's being fair to all students."*
AI Use Challenges / Opportunities	Personalized Learning Potential	Learners recognize that AI could provide tailored instruction or feedback based on individual weaknesses.	Lrn. 5: *"If it can notice I keep using incorrect verb tenses in reports, it should focus on correcting that next time."*
	Risks of Overdependence & Errors	Concerns about accuracy and critical thinking arise when learners rely too heavily on AI.	Lrn. 9: *"Sometimes the explanation is too generic or even wrong — especially in physics. I double-check it with my teacher."*
	Academic Integrity & Plagiarism	Some worry about students copying AI outputs without understanding, risking academic dishonesty.	Lrn. 6: *"My friend submitted a whole paragraph from ChatGPT and didn't even read it. That's dangerous."*
Conceptual Understanding of AI	AI Definition in Learning	Learners commonly perceive AI as an intelligent aid for writing, providing feedback, and offering clarification	Lrn. 1: *"To me, AI means something that helps you understand or generate academic English better, especially for complex concepts."*
	AI as a Learning Facilitator	AI supports learning by simplifying complex concepts, offering language suggestions, and providing ongoing feedback.	Lrn. 10: *"When I asked it to explain a marketing term, it gave both a definition and a sample sentence in business context."*
AI Design Expectations	Field-Specific Functionality	Learners desire features that cater to their disciplinary vocabulary and writing genres.	Lrn. 7: *"It should know medical writing tone—like patient cases or diagnostic reports—not just general essays."*
	Transparent, Justified Feedback	They want feedback to come with clear rationale, not vague corrections.	Lrn. 11: *"If it tells me to change a phrase, I want to know why—not just 'sounds better.'"*
	Institutional Policy & Training	Learners call for university-level AI use guidelines and training on ethical use.	Lrn. 1: *"I think we need orientation on how to use AI tools right, not just be told to avoid them."*

Table A5

Needs Identified from RICs (Workshop 1)

Theme	Frequency	Roles Mentioned	Illustrative Quote
Scaffolded Prompting Strategies	8	Instructors, Learners, AI Experts	*Prompts must be designed with an architect's eye to capture what AI misses.* (Ins.3)
Explicit Human-Check Mechanisms	9	Instructors, Policy-Makers	*Include a 'human-check gate' after AI output before submission.* (PM1)
Assessment Rubrics Fostering Engagement	7	Instructors, Learners	*Criteria should emphasize learner cognitive processes, not just surface fluency.* (Ins.2)
Transparent AI Decision Paths	5	AI Experts, Learners	*Explain why the AI gives a certain suggestion, not just the suggestion itself.* (Exp1)
Context-Specific Resource Repositories	4	Instructors, Policy-Makers	*We need resources that reflect Iranian ESP contexts.* (PM2)
Adaptive Feedback Loops	3	Learners, AI Experts	*Let the system re-check after I edit my work.* (Lrn.3)
Continuous Teacher Training	4	Policy-maker	*Our staff need annual refreshers on both the tech and the pedagogy.* (PM3)
Integration of Multimodal Inputs	3	Instructors, Learners	*I want to upload diagrams and have the AI integrate them into the text.* (Ins.3)
Real-time Error Flagging	3	AI Specialist	*System should detect and flag hallucinations before learner sees them.* (Exp. 3)

Table A6

Contextual Concerns Identified from RICs (Workshop 1)

Theme	Frequency	Roles Mentioned	Illustrative Quote
Over-Reliance on AI Outputs	8	Instructors, Policy-Makers, Learners	*Students might stop thinking critically if AI always answers for them.* (Ins3)
Bias in AI Training Data	6	AI Experts, Instructors	*The AI keeps giving Western examples that don't fit here.* (Exp.2)
Infrastructure Limitations	6	Policy-Makers, Learners	*Our lab computers crash when running AI tools.* (Lrn.1)
Data Privacy and Security Risks	4	Policy-Makers, AI Experts	*We cannot let sensitive documents leave our servers.* (PM2)
Language Drift from ESP Terminology	3	Instructors, Learners	*AI sometimes uses general English where technical terms are needed.* (Ins.1)
Erosion of Instructor Authority	2	Instructors	*If AI answers everything, why am I here?* (Ins2)
Ethical Ambiguity in Feedback	1	Learners	*The AI sometimes says things that sound disrespectful.* (Lrn. 3)

Table A7

Suggestions from RICs (Workshop 1)

Theme	F	Roles Mentioned	Illustrative Quote
Human-Check Gate Implementation	6	Policy Makers, Instructors, AI Experts	Establish a formal verification stage to review AI outputs before learner submission to ensure accuracy and ethical integrity.
Scaffolded Prompt Templates	5	Instructors, Learners	Develop discipline-specific prompt frameworks that support precise and relevant AI interactions.
Continuous Teacher Training	4	Policy Makers, Instructors	Provide ongoing professional development for educators focusing on AI literacy and ethical use.
Transparent AI Decision Paths	3	AI Experts, Instructors	Enhance the explainability of AI's decision-making processes to build trust among users.
Context-Specific Resource Repositories	3	Learners, Instructors	Curate specialized resources and materials tailored to ESP to better assist AI tutoring and learning scaffold.
Adaptive Feedback Mechanisms	2	Learners, Instructors	Implement responsive feedback loops personalized to individual learner needs and progress.
Data Privacy Protocols	2	Policy Makers, Learners	Establish strong institutional policies to safeguard learner data privacy within AI platforms.