Chapter 04 / Capítulo 04

New literacies in the age of AI: Ethics, teaching, and writing (English Version)
ISBN: 978-9915-9854-5-9
DOI: 10.62486/978-9915-9854-5-9.ch04
Pages: 45-61

Generative AI-assisted teaching strategies for designing cost structures with rigor and scalability

Estrategias didácticas asistidas por IA generativa (IA Gen) para diseñar estructuras de costos con rigor y escalabilidad

Car-Emyr Suescum Coelho[1], Carluys Suescum Coelho[2], Carlysmar Suescum Coelho[3]

[1]Universidad Metropolitana, Universidad Central de Venezuela. Caracas, Venezuela.

[2]Centro de Estudios Gerenciales Avanzados (CEGA). Caracas, Venezuela.

Corresponding author: Car-Emyr Suescum Coelho

ABSTRACT

This chapter examines how to rigorously and scalably integrate generative artificial intelligence into the teaching of cost structures in higher education. The objective is to design a teaching framework and strategy architecture with *prompts* aimed at strengthening accounting reasoning, traceability, and formative assessment. An evidence-based instructional design approach is adopted that articulates active learning, supervised algorithmic tutoring, and verification protocols. Additionally, prototypes are proposed for activity-based costing, time-based extensions, and cost-volume-profit analysis, along with prompt engineering rubrics and learning outcome matrices. The findings point to significant improvements in process transparency, the relevance of assumptions, and the reasoned defense of decisions when these are based on the documentation of interactions, triangulating product, process, and calculation. It is also noted that quality depends on data curation, critical literacy, and meaningful human oversight. The implementation of responsible use policies, the creation of curriculum-aligned prompt libraries, and the development of assessment cycles with iterative feedback are recommended. The contribution lies in translating the current discussion on generative AI into operational practices that enhance the quality of learning (SDG4) and the reliability of accounting decisions.

Keywords: Generative Artificial Intelligence; Cost Accounting; Prompt Engineering; Activity-Based Costing; Formative Assessment.

RESUMEN

El capítulo examinará cómo integrar con rigurosidad y escalabilidad inteligencia artificial generativa en la enseñanza de estructuras de costos en educación superior. El objetivo parte del diseño de un marco didáctico y una arquitectura de estrategias con *prompts* orientados al fortalecimiento del razonamiento contable, la trazabilidad y la evaluación formativa. Se adopta un enfoque de diseño instruccional basado en evidencias que articulan el aprendizaje activo, la tutoría algorítmica supervisada y los protocolos de verificación. Aunado a ello, se proponen prototipos aplicados a costeo basado en actividades, extensiones basadas en tiempo y análisis costo – volumen - utilidad, junto con rúbricas de ingeniería de *prompts* y matrices de resultados de aprendizaje. Los hallazgos señalan mejoras significativas en la transparencia del proceso, en la pertinencia de los supuestos y en la defensa argumentada de decisiones cuando éstas forman parte de la documentación de las interacciones, triangulando producto, proceso y cálculo. Se advierte además que la calidad depende de la curaduría presente en los datos, la alfabetización crítica y el control humano significativo. Se recomienda la implantación de políticas de uso responsable, la generación de bibliotecas de *prompts* alineadas al currículo y el desarrollo de ciclos de evaluación con retroalimentación iterativa. El aporte radica en traducir la discusión actual sobre IA generativa a prácticas operativas que eleven la calidad del aprendizaje (ODS4) y la confiabilidad de la decisión contable.

Palabras clave: Inteligencia Artificial Generativa; Contabilidad de Costos; Ingeniería De *Prompts*; Costeo Basado En Actividades; Evaluación Formativa.

INTRODUCTION

Teaching cost structures faces a sustained challenge: translating the complexity of real-world processes into the classroom without losing conceptual clarity or rigor in measurement. This is where the emergence of generative artificial intelligence (Gen AI) reconfigures this challenge by offering conversational tutoring, example generation, and immediate feedback, while also demanding critical literacy, verifiability, and authorship responsibility. In this context, the chapter aims to design and justify an AI-assisted teaching ecosystem for learning ABC costing, time-based extensions, and cost-volume-profit analysis with high standards of quality, ethics, and scalability.

García Peñalvo et al. (2024) frame the opportunities and challenges of this transition, arguing that generative AI, when integrated for explicit pedagogical purposes, can strengthen accounting reasoning by making visible the process of model construction, the choice of drivers, and the evaluation of scenarios.

This primarily requires documenting prompts, outputs, and decisions and, above all, submitting the calculation for verification with auditable traces. The quality of the interaction depends on competence in prompt engineering, which acts as a mediator between learning objectives, case study data, and output format, and must be evaluated with valid and reliable instruments (Gutiérrez Rosado et al., 2025).

The chapter is structured in four sections or guiding axes. First, a didactic framework is presented that articulates active learning, artificial intelligence (AI)-guided tutoring, and critical literacy, with operational criteria of privacy, verification, and human control. Second, an architecture is proposed that articulates strategies and prompt templates to orchestrate roles, flows, and iterations in cost tasks. Third, prototypes for ABC, TDABC, and CVU are described that perfectly exemplify how AI amplifies exploration and justification, and the criteria of rigor, evaluation, and scalability are finally discussed, which allow for the institutionalization of quality practices (Alier-Forment et al., 2025).

The relevance of this approach is supported by disciplinary evidence, as the literature on cost accounting indicates that the activity-based approach improves the allocation of inputs, the analysis of processes, and managerial decision-making when the binding causal relationship between resource consumption and cost objects is preserved (Casanova Villalba et al., 2021). This is precisely where generative AI enables rapid testing and iteration of scenarios and sensitivities without sacrificing the reasoned defense of assumptions or the reconciliation of totals, promoting the articulation of transferable learning and more informed decisions in commercial, industrial, and service contexts.

The chapter seeks to translate the general discussion on AI in higher education into an operational design for teaching cost structures, with well-defined responsibilities, robust verification protocols, and process-focused assessment. For Baldrich & Domínguez-Oller (2024), this avoids technocratic dependence and aims to achieve a level of literacy that combines an exceptional triad: technical precision, operational ethics, and disciplinary communication.

Given these precisions, academia is invited to consider the use and implementation of generative AI as a tool to make learning more transparent and accounting decision-making more robust, without ever renouncing the requirement to argue, analyze, calculate, and verify with professional judgment.

DEVELOPMENT

Didactic framework and principles for integrating generative AI into costs

The integration of generative AI into cost structure courses in our academic centers, from a didactic point of view, requires the articulation of learning that, beyond being active, is guided by tutorials mediated by artificial intelligence agents (AI agents) and a continuous evaluation process. However, critical literacy in AI is necessary to understand not only its much-praised capabilities but also its implicit limitations and risks, and to support solid institutional policies that guide its use from a pedagogical perspective.

Technology, and above all these tools, must be at the service of the judgment and decision-making capacity of the accounting professional, reinforcing not only the autonomy of the student, but also, as pointed out by García Peñalvo et al. (2024), evaluative transparency, which stems from the need to train teachers and students and goes hand in hand with the revision of curricula to ensure that its use is profoundly ethical and humanistic.

Pedagogically speaking, García Sánchez (2023) offers a clear synthesis of these fundamentals when he argues that generative AI is articulated with behavioral, cognitive, constructivist, and even connectivist approaches. However, this is precisely why its implementation and relevance lie in the design of meaningful experiences that connect immediate feedback with active meaning construction.

It is at this point that the role of virtual tutors comes to the fore, facilitating metacognitive scaffolding and simulations in which decisions must be made about costs, while networked platforms promote collaborative learning and evidence curation. At this point, the diversity of paradigms may require explicitly stating basic assumptions and quality criteria before delegating task performance to a model, thereby transforming early AI literacy into epistemological literacy.

Tutoring guided by artificial intelligence tools takes on real meaning when it is integrated symbiotically with practical, real-life case designs that require in-depth, structured cost analysis and professional reporting. Given this requirement, authors such as Alier-Forment et al. (2025) recommend conversational assistants incorporated into the course corpus to ensure the traceability of sources and the orientation of specific tasks, as is the case with the linking of questions and concerns that students generally present with certain course content and audiovisual material, thus strengthening validation, active learning, investigative autonomy, and the coherence of their final results.

As has been the case throughout human history, for any adoption to be considered truly responsible, a set of explicit principles must be guaranteed, such as privacy, protection of student confidentiality, algorithmic transparency, strategic alignment with the curriculum, development of transparent interfaces, and adequate teacher supervision.

To achieve this, it is also necessary to minimize the cognitive load introduced by complex tools and to ensure verification mechanisms that prevent the acceptance of speculative answers as valid. Alier-Forment et al. (2025) formulate practical guidelines throughout their research that guide this pedagogical governance of generative AI in university contexts, where the operational criteria are ultimately translated into rubrics and protocols that serve as accompaniment for each activity.

From this perspective, the mastery required in the discipline requires anchoring the framework in results specific to learning cost accounting, namely, modeling structures, analyzing situations, and the fundamentals and arguments of decision-making. From this perspective, when deploying courses with generative AI, it is necessary to require, in the words of Casanova Villalba et al. (2021), the balances and perspectives that justify this emphasis on structural cost analysis, that is, the student must argue with evidence their choice of allocation bases, the identification of activities involved, and the articulation between costs and processes, since the aim is not to automate reasoning but to strengthen it with the acquisition of fundamentals, verification, and traceability.

Da Costa Marques (2012) substantiates the contribution of ABC in university contexts, as this approach provides a robust basis for translating resource consumption into activity and product costs, a central aspect in designing prompts for the simulation of reallocation scenarios. Improving the quality of indirect cost information will directly favor process management, making it ideal for AI-guided tasks that require triangulation and critical review.

The practical relevance of the framework is strengthened by sectoral evidence obtained, for example, in logistics, where the costs associated with activities can become the second most important component after production. This situation requires modeling processes, times, and critical points before making decisions. This is where the argument put forward by Zúñiga Marín & Aguirre González (2022) becomes most relevant, when they state that generative AI supports mapping and sensitivity analysis, provided that the data and assumptions are auditable.

Segarra Ciprés et al. (2024) report these tensions and opportunities in their teaching experience in the higher education system. The authors indicate that AI literacy also involves understanding teachers' perceptions, concerns, and risks, which range from the veracity of responses to the confusion that may arise between the use of the tool and learning achievement. However, beyond latent concerns, there are recognized benefits to streamlining tasks and exploring topics, as competency mapping must integrate not only verification skills but also critical reasoning and the ethical treatment of sources, along with strategies that enable self-regulation of learning.

The development of prompt engineering competency thus becomes a cross-cutting axis for articulating disciplinary rigor with metacognition. In this regard, Gutiérrez Rosado et al. (2025) highlight the need to evaluate this competency in higher education validly and reliably. Therefore, the rubric must consider intentionality, context, restrictions, success criteria, and the traceability of sources, in addition to reflecting on the algorithmic biases and hallucinations presented by LLMs. Likewise, the teaching ecosystem must integrate microcycles of design, testing, and prompt review, with evidence commented on in the respective portfolios.

Suárez-Martínez et al. (2025), throughout their research, show an evaluative method based on prompt engineering with augmented retrieval, highlighting that the evaluation of media products and artifacts generated by AI must be based at all times on clear indicators and explainable analytics that effectively capture the quality, learning, and transmedia production when applicable (Sarmiento Contreras & Zamora Arízaga, 2025). In addition, it is relevant to use frameworks such as the GHQ (Goal Question Metric) to align objectives, questions, and metrics, always with transparency regarding training data and review procedures (2008). Thus, the rubrics used in courses should incorporate evidence of prompt iteration, along with corresponding justifications for decisions.

The coherence of the framework is supported by recent evidence reporting improvements in personalization, participation, and immediate feedback when generative AI is integrated into adaptive designs. Romani Pillpe et al. (2025) synthesize these findings and warn of the importance of AI literacy for instructional design, further emphasizing that such benefits require deeply ethical literacy, solid corpus curation, and the implementation of verification protocols to avoid accountable decisions based on opaque outputs, thus consolidating all the criteria of relevance, accuracy, traceability, and equity that are necessary to reform the critical analytical capacities of professional judgment.

Strategy, architecture, and prompt ecosystem

The design of a prompt ecosystem for courses that teach cost structures requires an architecture that connects learning objectives, accounting data, and pedagogical verification. It is therefore necessary to propose a procedural sequence that integrates teaching planning, conversational tutoring, and quality control, each with traceable evidence of the process (Torres & Blanco, 2023).

In addition, technological orchestration must be aligned with institutional policies and a process of progressive AI literacy, so that the tool complements and exponentially enhances professional judgment rather than replacing it. Such curricular decisions require clear frameworks for evaluation, ethics, and teacher training (García Peñalvo et al., 2024).

This is where the architecture organizes roles and flows. At the same time, the teacher designs cases, defines criteria, and validates assumptions, which is reinforced by the conversational tutor guide (AI agent), which exemplifies and suggests alternatives, using an automated verifier to check internal consistency and reconstruct calculations, so that, ultimately, the student argues and defends their solution (Torres Vargas, 2024). Consequently, Borja Borja (2025) demonstrates that personalized tutoring with AI can increase autonomy when there is a precise pedagogical design, as the interaction becomes an iterative cycle of question, answer, and refinement, with explicit goals and public rubrics, favoring personalization, motivation, and self-regulation, as long as the tutor is used as scaffolding and not as a shortcut (Suescum Coelho et al., 2025).

According to Gutiérrez Rosado et al. (2025), the technical core of the ecosystem comprises prompt templates, task contracts that specify a purpose, context, data, and the corresponding output format. This is where the quality of the prompt is evaluated first for clarity, then for completeness, data relevance, success criteria, and, above all, its traceability. It is therefore a competent approach that, by incorporating metacognition, requires justification of assumptions, forces detailed verification, and facilitates the reliable measurement of these dimensions.

For Puerto & Ruiz (20254), the operational sequence includes a pre-diagnosis that allows prior knowledge to be activated, that is, an initial prompt to generate an initial solution and a kind of guided re-prompt that ultimately improves accuracy, explainability, and consistency; the output is not reduced to the simple final result, but documents iterations and ongoing decisions.

Meza Arguello et al. (2025) point out that evidence in higher education suggests that the perceived usefulness of assistants increases when teachers teach them to refine questions and demand more auditable output formats, increasing the transparency of the learning process, which translates into significant improvements in understanding when working with guided iterations and verifiable products (Segarra Ciprés et al., 2024).

In the area of costing, prompt templates should prompt the creation of an activity dictionary, the identification of drivers, the estimation of rates, and the reconciliation between resource costs and allocated costs. This is where an effective prompt will specify the cost object, delimit processes, provide structured data, and request tables with calculation traces. For Da Costa Marques (2012), the emphasis lies precisely on explaining why an inducer is causal and how the cost varies under different scenarios. Only in this way can the quality of the decision be improved by translating resource consumption into information relevant to management (Juca et al., 2024).

For Albarracín Vanoy (2023), in logistics environments, it is helpful to have a set of templates that connect time, capacity, and transfer points to the allocation of indirect costs. In this regard, Zúñiga Marín & Aguirre González (2022) illustrate how a well-parameterized ABC allows for the identification of sources of inefficiency in supply chains, recommending that assumptions about volume, variability, and seasonality be made explicit, and, in addition, requiring sensitivity analysis with limited percentage changes, which will verify the duty to check the closure of sums and consistency between activities.

Cost-volume-profit analysis requires templates designed to structure margins, break-even points, and scenarios with simultaneous variations in price, mix, and fixed costs. That is why, at this point, the task contract must require the reporting not only of assumptions, but also the derivation of formulas and the tabular presentation of multiple scenarios, accompanied at all times by a reflection on sensitivity and risk. Casanova Villalba et al. (2021) emphasize that pedagogical value truly emerges when quantification is connected to narratives of management and continuous improvement.

The architecture outlined in the prompt requires, in Zapata Ros's (2024) words, a quality control module that translates the evaluation into verifiable evidence. In this vein, he proposes that each delivery include the version of the prompts, the assistant's output, and a human verification annex that recalculates and justifies discrepancies. This is how a verification agent can be deployed to review these tables, identify closings, and flag inconsistencies. Under this operational logic, the reliability of the process is reinforced, and co-evaluation is facilitated, a co-evaluation that not only favors replicability but also the process of meta-reflection (Suárez-Martínez et al., 2025).

Baldrich & Domínguez Oller (2024) argue that the writing dimension of the ecosystem requires transparency about AI's contributions in the corresponding reports and technical notes. Therefore, they maintain that each template used should include at least three sections: responsible authorship, citation criteria, and a statement of the assistant's limitations. In addition, it is advisable to promote cross-peer review strategies to strengthen style, consistency, and the use of evidence. Research on academic writing with conversational assistants (González, 2025) indicates gains when teaching guidance, peer review, and a strong emphasis on disciplinary argumentation are combined, without delegating epistemological responsibility.

Sustainable adoption of the ecosystem benefits from a realistic framework that not only addresses expectations but also reinforces them with training support that addresses concerns, perceptions, doubts, and risks of dependency. The architecture must, prima facie, incorporate an initial orientation, teaching contracts that are sufficiently explicit, and, above all, support mechanisms so that students with less technological familiarity are not left behind. García Sánchez (2023) notes in his research that mixed attitudes become more favorable when purposes are clarified, when guided examples are offered, and, above all, when standards of academic honesty are established, emphasizing the importance of transparency and accountability to maintain trust in evaluation processes.

The external consistency that must prevail in the ecosystem rests on the convergence of studies showing benefits from personalization, feedback, and engagement when AI is effectively integrated, with clear criteria for systematic verification (Villacreses et al., 2025). Recent reviews, such as that developed by Romani Pillpe et al. (2025), recommend moving toward evidence-based designs, with metrics aligned with learning outcomes and protocols that reduce the opacity of models. An iterative and documented approach is therefore required to facilitate educational quality assurance and scale successful experiences not only in terms of costs but also in different subjects.

Governance must be based on ethical and operational principles that protect privacy, academic integrity, and, above all, the maintenance of meaningful human control, primarily because the architecture of these prompts must limit access to or violation of sensitive data, promoting the approval and requirement of understandable explanations that guarantee that the final decision remains in human hands, that is, with the student and the teacher as facilitators of the process. It is advisable to institutionalize rubrics, protocols, and repositories of validated templates in this way, so that they evolve with practice, balancing innovation, quality, rigor, and educational responsibility. (Alier-Forment et al., 2025).

Learning prototypes for cost structures

This section describes educational prototypes that integrate generative AI to support learning, practice, and evaluation of the design of cost structures in commercial, industrial, and service contexts. The emphasis is on pedagogical replicability and progressive improvement through prompt engineering, so that students progress from guided tasks to autonomous and transferable performance. In this sense, it is assumed that competence in prompt design requires explicit, assessable criteria, as a flexible, adaptable rubric maintains the validity of the process in the face of rapid technological evolution and helps overcome approaches that focus solely on the final product. Gutiérrez Rosado et al. (2025) support this need and offer methodological bases for its implementation in higher education.

The prototype focuses on building an ABC model from a dataset containing indirect costs, activities, and cost objects. AI is used to suggest initial activity schemes and propose plausible drivers, which the student contrasts with evidence from the case. Consequently, a refinement cycle is promoted in which ambiguities are corrected, activities are relabeled, and drivers are justified with causality criteria. This itinerary converges with the canonical definition of ABC as a two-stage measurement and allocation method, proper not only for calculating costs but also for supporting management control and strategic decision-making. Da Costa Marques (2012) explains this by detailing the transition from resources to activities and from activities to cost objects.

As can be seen, algorithmic support is particularly fruitful when the case involves extensive activity chains, product heterogeneity, or scale variations. In this vein, a second prototype applies the ABC scheme to logistics processes, asking AI to help map key activities, estimate relative consumption, and highlight cause-and-effect relationships between drivers and tasks. A student contrasts each suggestion with observed data and contextual assumptions, reducing bias and verifying consistency.

Zúñiga Marín & Aguirre González (2022) document the relevance of this approach by emphasizing the importance of causal drivers and measuring activity consumption to obtain more reliable costs in profitability decisions. On the other hand, a third prototype is proposed to extend the analysis to standard time-based schemes, where the student formulates time equations for activities and uses AI to test scenarios of capacity, bottlenecks, and alternative allocations.

The model's function is not to replace professional judgment under any circumstances, but rather to accelerate the identification of critical assumptions and their sensitivity to variations in scale or product mix. In this way, students learn to integrate operational measurements into a cost system, with the dual purpose of achieving informational accuracy alongside decision-making discipline. Casanova Villalba et al. (2021) emphasize that rigorous cost management, supported by methods that detail items and activities, strengthens decision-making and aligns information to generate truly sustainable profits.

In addition, a fourth prototype focused on cost-volume-profit analysis is proposed. AI acts as a co-producer of scenarios by generating banks of assumptions about prices, elasticities, sales mix, and fixed and variable cost structures. At the same time, the student validates the consistency and relevance of each case. Therefore, the practice combines financial interpretation with hypothesis testing, incorporating sensitivity analysis and discussion of safety margins.

This makes it clear that the educational goal is for students to move from mere exercises that, at first glance, appear algorithmic to deeper narratives of more profound decision-making, where they can justify break-even thresholds, assess the impacts of the mix, and consider alternatives for improvement that benefit the organization. It is within this framework that cost management is inseparable from planning and evaluating results, because, as Casanova Villalba et al. (2021) argue, highlighting cost information accurately strengthens the balance between operational decisions and profitability objectives.

In prototypes, generative AI plays a tutorial role, guiding exploration step by step, offering immediate feedback, and adapting the level of scaffolding based on observed performance. For this reason, Posso-Pacheco (2025) argues that this type of interaction promotes autonomy without sacrificing methodological rigor, since each AI suggestion will require verification not only with data, but also with accounting principles.

Borja Borja (2025) emphasizes that it is precisely this personalized support that enhances retention and self-management of learning in higher education, given that algorithmic tutoring is integrated and mimics teaching mediation, reinforcing content, motivation, and self-regulation when there is clear pedagogical guidance.

The effectiveness of prototyping depends mainly on sound prompt engineering. Torrealba Dugarte (2024) recommends starting with clear contextual instructions that establish the role, the objective pursued, the output format, and the quality criteria; then iterating with requests for cross-checking, searching for counterexamples, and explanations of critical assumptions. It is through the development of this pattern that the accuracy of responses and the structure of sequential reasoning are improved, as the student, step by step, avoids giving generic answers, thereby promoting greater didactic traceability. Entering clear prompts with explicit context and constraints guides the model and yields much more valuable and relevant results in educational interaction (Segarra Ciprés et al., 2024).

The evaluation of prototypes should include a meta-communicative layer that documents objectives, representative tasks, validation of the responses generated by the system, but primarily the iterative redesign of interactions, as this is relevant in the articulation of communicative evaluability methods with the adaptive interaction factors of the prompts, so that the student themselves learns to audit and detect possible biases, errors, and deviations, taking on a much more active role and allowing them to propose improvements. Suárez-Martínez et al. (2025) describe this methodological harmonization by outlining how to make explicit the stages of preparation, implementation, validation, and feedback to ensure coherence between pedagogical intention and the results generated by AI. However, there are caveats, and adoption is not without precautions that must be considered, as evidence suggests heterogeneous student perceptions of the accuracy, usefulness, and teachers' preparedness to integrate these tools into the classroom (Mosquera-Gende & Canut, 2025). These perceptions reinforce the importance of making criteria explicit, measuring learning, and maintaining critical literacy in AI that enables responsible use. In specific educational contexts, there are disparate assessments of its suitability and accuracy, as well as doubts about the preparation teachers have received for its incorporation. This situation, in effect, requires institutional policies and continuous training (García Sánchez, 2023).

The prototypes described articulate cost modeling, algorithmic tutoring, and metacommunicative assessment to develop technical and metacognitive skills with traceability. Consequently, the emerging training ecosystem not only accelerates mastery of ABC, time-based extensions, and CVU analysis but also enables personalization and student engagement when designed with transparency and ethics. Based on this, Romani Pillpe et al. (2025) conclude that integrating generative AI into adaptive and personalized environments increases student participation, provided literacy and ethical considerations are addressed.

Rigor, evaluation, and scalability of the ecosystem

For García-Peñalvo (2024), the rigor that should govern generative AI-assisted learning ecosystems must be based on verifiable criteria, transparent procedures, and traceable evidence of learning. Therefore, assessment cannot be reduced to the mere mechanical or quasi-automated detection of machine-generated texts, since such detectors exhibit empirical limitations that compromise fair academic decisions (Franganillo, 2023).

It is under these considerations that it is more prudent to shift the proper focus to performance quality, to the explicitation of processes, and, above all, to consistency with the learning outcomes of the cost structures course, as this is how evaluative fairness and reputational security for students can be ensured, according to recent literature on AI in education (López De La Cruz, 2024).

Suárez-Martínez et al. (2025) point out that assessment should triangulate evidence: the final product of the task, the process of interaction with AI, and the student's own accounting reasoning. In addition, it should incorporate traces of the prompts used during the session, brief justifications, and references to the subject matter, allowing it to assess how algorithmic assistance is transformed into meaningful learning. This is where the approach is consistent with assessment experiences mediated by prompt engineering for the critical appraisal of these educational artifacts, using the simulation of expert panels with generative AI to contrast criteria, thus enriching formative feedback.

Rigor requires recognizing prompt engineering as an assessable, cross-cutting, and transferable competency. Therefore, a specific instrument, such as the prompt engineering rubric validated for higher education (Gutiérrez Rosado et al., 2025), provides performance descriptors and achievement levels that will ensure inter-evaluative consistency and clarity for students. Its use organizes the observation of key skills, from the specification of the objective to responsible iteration, and promotes comparability between cohorts, causing the evaluation to shift its focus from isolated results to capturing the quality of the inquiry process.

A professionally oriented ecosystem must anchor assessment to disciplinary practices. In cost structure courses, this involves assessing activity identification, driver selection, and indirect cost allocation, in line with ABC approaches that provide more rational information for decision-making. On the other hand, the design of AI-guided tasks should require students to make assumptions explicit, validate data, and compare scenarios, so that automation complements, rather than replaces, accounting judgment (Da Costa Marques, 2012).

Likewise, the empirical literature on ABC in operational contexts shows that accuracy in activity modeling affects cost reduction and strategic planning. Therefore, IAG assessment may ask students to construct activity maps, estimate resource consumption, and compare actual distributions with optimal ranges, recording their verification of each assumption in prompts and reflection notes. This articulation allows for judging not only the correctness of the model, but also the quality of AI-supported decision-making (Zúñiga Marín & Aguirre González, 2022).

Rigor also demands clear and socialized operational ethics. Some classroom experiences with ChatGPT show benefits in discursive cohesion and argumentative structure, but also warn of risks of dependency, diffuse authorship, and plagiarism if its use is not guided correctly. Baldrich & Domínguez-Oller (2024) argue that assessment should include criteria for integrity, citation, and intellectual originality, as well as spaces for meta-reflection on the tool's limits and scope. Therefore, making these criteria explicit in the guidelines and rubrics would avoid misunderstandings, strengthening student autonomy and responsibility.

The ecosystem's scalability is supported by architectures that integrate assistants aligned with the curriculum and verified sources, capable of providing clues, examples, and contextual references. When the assistant documents its responses with citations and excerpts from course materials, the teacher gains traceability, and the student receives relevant, auditable support. This facilitates replication across subjects and cohorts and maintains a standard of pedagogical quality even with large groups. Alier-Forment et al. (2025).

At the same time, comparative evidence on IAG in education suggests advantages in the design of experiences and large-scale support, provided that disciplinary relevance and pertinence are taken into account. In addition, the findings indicate that AI literacy modulates the benefit: those who master contextual and collaborative prompts obtain better results than those who use generic prompts. Therefore, the ecosystem must also evaluate the ability to adapt prompts, combine sources, and review biases as part of learning outcomes (Romani Pillpe et al., 2025).

From a course management perspective, it is advisable to establish formative assessment cycles that combine micro-deliveries with feedback from AI and the teacher. In this sense, algorithmic clues become inputs for human review, not substitutes. In addition, student perceptions of benefits and ethical dilemmas should be monitored with brief instruments, as they influence the acceptance of strategies and their effectiveness. This pedagogical monitoring, with usage data and qualitative evidence, enables transparent adjustments to criteria and workloads (García Sánchez, 2023).

To consolidate validity and reliability, it is recommended to align performance rubrics with learning outcome matrices and cost domain descriptors. For example, the assessment of an ABC prototype with IAG could consider the relevance of activities, the consistency of drivers, the justification of assumptions, and the quality of sensitivity analysis. Complementarily, the prompts rubric would judge the clarity of the objective, the use of context, and iteration decisions. As can be seen, both dimensions converge on triangulated, replicable evidence across semesters. Casanova Villalba et al. (2021).

On the other hand, the reliability of the assessment requires caution when it comes to technological solutions that promise to detect generated texts automatically. Research shows that such detectors are not accurate enough for high-impact decisions, so their use should be avoided as conclusive evidence. In this context, evaluating reasoning, decisions, and work traces is methodologically more robust and reduces false positives that would unfairly affect student performance. Therefore, assessment policy should make this criterion explicit and offer alternative means of verification (García Peñalvo et al., 2024).

Scalability is also achieved through prompt libraries aligned with the curriculum and prompt engineering guidelines, enabling different teachers to apply a common standard. These libraries should be structured by cost scenarios, activities, and constraints, facilitating transfer between subjects and consistency in the student learning experience. In addition, prompting training itself must be grounded in the principles of clarity, context, and guided iteration, as described by Segarra Ciprés et al. (2024) in the university setting.

A decisive component of scalability is personalized tutoring supported by IAG, which can offer differentiated scaffolding without overwhelming the teacher. At this point, if the assistant incorporates validated knowledge bases and generates traceable feedback, teaching benefits from constant and consistent support, and the student progresses with greater autonomy (Borja Borja, 2025). For these reasons, assessment must recognize not only these personalized itineraries, but also their impact on the achievement of competencies, sequentially documenting evidence of both the process and the product.

Methodological rigor requires protocols for continuous review and improvement of the ecosystem. García Peñalvo et al. (2024) argue that, in addition to analyzing achievement rates and product quality, it is advisable to audit algorithmic biases, the relevance of sources, and the adequacy of cost drivers, and to incorporate progressive adjustments. It is suggested that the combination of rapid literature reviews, case studies, and experience evaluation offers an effective route to sustaining quality and relevance, even in scenarios of rapid change driven by the evolution of generative AI.

For the ecosystem to be rigorous, assessable, and scalable, it must articulate three pillars: competencies in prompt engineering with explicit, validated criteria; disciplinary tasks that require accounting judgment in cost structures; and assistance architectures with traceability and operational ethics. On the other hand, prudence in the use of detection tools and the centrality of disciplinary reasoning will be the pillars that sustain academic equity. Therefore, when assessment considers both performance and process, and when AI devices are embedded in robust pedagogical frameworks, scalability ceases to be an abstract promise and becomes a verifiable, sustainable practice (Romani Pillpe et al., 2025).

FINAL REFLECTIONS

The incorporation of generative artificial intelligence into the teaching of cost structures has proven to be an opportunity to strengthen accounting reasoning, provided it is governed by criteria of quality, ethics, and traceability. In this sense, the technological resource does not replace the student's professional judgment or the teacher's mediation, but rather amplifies the ability to explore scenarios, contrast assumptions, and communicate decisions in a reasoned manner. Recent evidence suggests that this integration requires a review of assessment practices and an update of curricular frameworks to ensure learning is relevant and equitable.

This chapter has argued that didactic rigor is expressed in the clarity of objectives, the quality of data, and the verifiable reconstruction of the calculation process. Therefore, it has insisted on documenting prompts, outputs, and decisions, and on requiring causal explanations for the choice of drivers and allocation bases. This logic is consistent with the nature of ABC and TDABC methods, where the passage from resources to activities and from activities to cost objects requires auditable traces and internal consistency.

Comparative experience suggests that a learning ecosystem with algorithmic tutoring can promote autonomy, self-regulation, and meaningful learning if the scaffolding avoids directive responses and encourages reflection. Therefore, it is advisable to prioritize assistants who ask questions, re-ask questions, and offer alternatives, rather than dictating closed solutions. This perspective shifts the focus from simple procedural correction to conceptual understanding and transfer to real cost management contexts.

It is important that prompt engineering be understood as a cross-cutting and assessable skill, as it mediates the quality of interaction and the depth of reasoning. A student who states purpose, context, data, and success criteria, and who iterates with verification, learns to think with explicit language and controlled evidence. Such metacognitive learning not only reduces the risk of technological dependence but also increases the intelligibility of the process for peer co-evaluation and teacher feedback.

Practices with generative AI are not free of tensions. Doubts persist about accuracy, bias, and authorship responsibility, as well as heterogeneous perceptions among teachers and students regarding usefulness, fairness, and workload. In this vein, it is advisable to combine initial guidance, guided examples, and clear teaching contracts that define permitted uses, mandatory verification, and minimum traceability. Institutional transparency and accountability will therefore be responsible for strengthening trust and informed acceptance of the strategies.

The writing and communication dimension is another area for improvement. When AI support is combined with peer review processes and explicit criteria for responsible authorship, progress can be seen in argumentative coherence and evidence quality, without diluting the requirement for critical thinking. This is particularly relevant in cost reports, where the narrative must explain decisions, risks, and sensitivity, rather than simply presenting spreadsheets.

From a disciplinary perspective, the prototypes presented demonstrate that the link between activity modeling, cost allocation, and sensitivity analysis gains new power when verifiable algorithmic tutoring is integrated. However, the pedagogical value only materializes when practice requires justifying causality, reconciling totals, and defending assumptions against alternative scenarios. In logistics contexts, for example, this requirement allows for the identification of inefficiencies and the reconfiguration of processes with greater precision and timeliness.

The evaluation of the ecosystem must balance product and process. Here, the accuracy of the calculation is important, but even more important is the quality of the reasoning behind it and the consistency of the trajectory of interaction with AI. The methodological option of integrating communicative evaluability and prompt rubrics is appropriate for auditing decisions, detecting biases, and promoting iterative improvements in interaction. This strategy also greatly facilitates replication across cohorts and the consolidation of task libraries and templates with shared standards.

Generative AI is not a shortcut to learning cost accounting, but a means of making reasoning more explicit, the process more transparent, and evaluation more rigorous. The central question is no longer whether the tool is accurate, but rather how its use is governed to enhance learning and the quality of accounting decisions. The evidence indicates that benefits are realized when there are clear frameworks, critical literacy, and systematic verification, conditions that this chapter has translated into concrete instructional design proposals.

The field of study remains open to new research agendas examining differential effects by entry profiles, traces of interaction that predict learning, and impact metrics that harmoniously integrate rigorous criteria of technical, ethical, and communicative quality. Similarly, it is pertinent to examine how prompting and modeling skills evolve throughout the curriculum, and to investigate which combinations of human and algorithmic tutoring would maximize understanding of cost structures in complex contexts such as those of today. This provides the scientific community with a foundation for moving forward with caution and ambition, avoiding false dilemmas, and, above all, prioritizing meaningful learning to achieve truly high-quality education.

REFERENCES

Albarracín Vanoy., RJ. (2023). Educational Transformation: Optimization in the Teaching of Logistics in International Business through the Application of Artificial Intelligence in Higher Education Institutions. Salud, Ciencia y Tecnología. https://doi.org/10.56294/sctconf2023422

Alier-Forment, M., Casañ-Guerrero, M., Pereira, J., García-Peñalvo, F., y Llorens-Largo, F. (2025). Inteligencia artificial generativa y autonomía educativa: metáforas históricas y principios éticos para la transformación pedagógica. RIED-Revista. https://doi.org/10.5944/ried.45536

Baldrich, K., & Domínguez-Oller, J. (2024). El uso de ChatGPT en la escritura académica: Un estudio de caso en educación. Pixel-Bit. Revista De Medios Y Educación, 71, 141–157. https://doi.org/10.12795/pixelbit.103527

Borja Borja, H. (2025). Inteligencia artificial generativa como tutor personalizado en la educación superior. Innovarium International Journal, 3(1). https://bit.ly/4hv7cjY

Casanova Villalba, C., Núñez, R,. Navarrete, C. y Proaño, E. (2021). Gestión y costos de producción, balances y perspectivas. Revista de ciencias sociales, 27 (1). https://bit.ly/4hAJpzp

Da Costa Marques, M. (2012). Contribución del modelo ABC en la toma de decisiones: el caso universidades. Cuadernos de Contabilidad, 13 (33), 527-543. https://bit.ly/47qokmj

Franganillo, J. (2023). La inteligencia artificial generativa y su impacto en la creación de contenidos mediáticos. methaodos. Revista De Ciencias Sociales, 11(2), m231102a10. https://doi.org/10.17502/mrcs.v11i2.710

García-Peñalvo, F. J. (2024). Inteligencia artificial generativa y educación: Un análisis desde múltiples perspectivas. Education in the Knowledge Society (EKS), 25, e31942. https://doi.org/10.14201/eks.31942

García-Peñalvo, F., Llorens, F., y Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. RIED-Revista Iberoamericana de Educación a Distancia, 27(1), 9–39. https://doi.org/10.5944/ried.27.1.37716

García Sánchez, O. (2023). Uso y percepción de ChatGPT en la educación superior. Revista De Investigación En Tecnologías De La Información, 11(23), 98–107. https://doi.org/10.36825/RITI.11.23.009

González, A.(2025). Escritura académica aumentada por IA generativa. Trayectorias Universitarias, 11(20). https://doi.org/10.24215/24690090e184

Gutiérrez Rosado, J., Palacios Flores, K., Cajo Pullas, M., y Villota Oyarvide, W. (2025). “Prompt engineering” como competencia transversal clave: desarrollo y validación de una rúbrica para su evaluación en la educación superior. RECIAMUC, 9(3). https://doi.org/10.26820/reciamuc/9.(3).julio.2025.209-223

Juca, F., Carchi, k., & Rosales, C. (2024). Análisis de costos mediante inteligencia artificial: un estudio comparativo de plataformas emergentes. Revista Dilemas Contemporáneos. 1 (12). https://doi.org/10.46377/dilemas.v12i1.4327

Koziolek, H. (2008). Goal, Question, Metric. In: Eusgeld, I., Freiling, F.C., Reussner, R. (eds) Dependability Metrics. Lecture Notes in Computer Science, vol 4909. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-68947-8_6

López De La Cruz, E. C. I. (2024). Inteligencia artificial en la justicia: un análisis teórico de metodologías y accesibilidad. Revista jurídica Peruana Desafíos En Derecho, 1(2), 91-96. https://doi.org/10.37711/RJPDD.2024.1.2.5

Meza Arguello, D. M., Obando Burbano , M. de los Ángeles, Franco Valdez, J. L., & Simisterra Muñoz, J. M. (2025). El uso de la inteligencia artificial como recurso pedagógico en la educación superior: experiencias de los docentes. Sage Sphere International Journal, 2(2), 1-10. https://doi.org/10.63688/jrd8es89

Mosquera-Gende, I., & Canut, L. (2025). Percepción docente sobre el uso de la Inteligencia Artificial en el contexto educativo. EDUCA. Revista Internacional Para La Calidad Educativa, 5(2), 1-22. https://doi.org/10.55040/86gnwm87

Posso-Pacheco, R. J. (2025). El rol del docente en la era de la inteligencia artificial: De transmisor de contenidos a formador de habilidades y valores para el siglo XXI. MENTOR Revista

De investigación Educativa Y Deportiva, 4(11), 1–8. https://doi.org/10.56200/mried.v4i11.10185

Puerto, A., & Ruiz, N. (2025). Inteligencia artificial como asistente pedagógico virtual para el fortalecimiento de las capacidades ejecutivas en el entorno educativo. LÍNEA IMAGINARIA, 1(20). https://doi.org/10.56219/lneaimaginaria.v1i20.3707

Romani Pillpe, G., Macedo Inca, K., Soto Loza, G., Franco Guevara, A. y Ore Choque, M. (2025). Revisión sistemática de inteligencia artificial generativa (GenIA) para el diseño de experiencias de aprendizaje, 2020-2025. Revista Espacios, 46(3). https://doi.org/10.48082/espacios-a25v46n03p02

Sarmiento Contreras, P. F., & Zamora Arízaga, R. X. (2025). Transmedia narrative in the audiovisual production of university academic content in Ecuador. Runas. Journal of Education and Culture, 6(12), e250234. https://doi.org/10.46652/runas.v6i12.234

Segarra Ciprés, M., Grangel Seguer, R. y Belmonte Fernández, Ó. (2024). ChatGPT como herramienta de apoyo al aprendizaje en la educación superior: una experiencia docente. Tecnología, Ciencia y Educación, 28, 7-44. https://doi.org/10.51302/tce.2024.19083

Suescum Coelho, C., Suescum Coelho, C.-E., Suescum Coelho, C., Suescum Coelho, C., & Coelho Freitas, C. M. (2025). Rol de los Agentes de Inteligencia Artificial en la Promoción de una Educación de Calidad: Una Perspectiva Basada en los Objetivos de Desarrollo Sostenible de la Organización de Naciones Unidas. Revista Científica Multidisciplinar SAGA, 2(2), 90-101. https://doi.org/10.63415/saga.v2i2.81

Suárez-Martínez, J., Arnau Roselló, R y Nieto-González, R. (2025). Evaluación de documental educativo con IAG utilizando prompt engineering. Comunicar, 33(82), 90–102. https://doi.org/10.5281/zenodo.16121533

Torrealba Dugarte, J. D. . (2024). Eduética en el uso de la inteligencia artificial a través de la ingeniería de *prompts*. Revista Digital De Investigación Y Postgrado, 5(10), 261-266. https://doi.org/10.59654/7j69c731

Torres, K., & Blanco, I. (2023). Arquitectura de prompt académicos para el uso de inteligencias artificiales (AI) en Areandina. Revista Investigaciones Andina, 25(46). https://doi.org/10.33132/01248146.2258

Torres Vargas, J. (2024). La ingeniería del prompt aliada de la inteligencia artificial. TRASCENDERE,5(1). https://doi.org/10.56219/trascendere.v5i1.2987

Villacreses, E., Moreira, N.,. C., Calderón, J., Torres, V., Iza, M., Tandazo, F., & Bernal, A. (2025). Inteligencia Artificial: Transformando la Escritura Académica y Creativa en la Era del Aprendizaje Significativo. Revista Científica De Salud Y Desarrollo Humano , 6(1), 1427–1451. https://doi.org/10.61368/r.s.d.h.v6i1.533

Zapata Ros, M. (2024). IA generativa y ChatGPT en Educación: Un reto para la evaluación y ¿una nueva pedagogía?. REVISTA PARAGUAYA DE EDUCACIÓN A DISTANCIA (REPED), 5(1), 12–44. https://doi.org/10.56152/reped2024-vol5num1-art2

Zúñiga Marín, J. & Aguirre González, E. (2022). Diseño de un modelo de costos basado en actividades aplicado a procesos logísticos. Caso: empresa del sector alimenticio tradicional. Revista EIA, 19(37), 1–18. https://doi.org/10.24050/reia.v19i37.1512

CONFLICT OF INTEREST

There are no conflicts of interest of any kind.

FUNDING

The work has not received any subsidies from public, commercial, or non-profit funding agencies.

USE OF ARTIFICIAL INTELLIGENCE

AI was used for the translation of the abstract and style correction.

AUTHORSHIP CONTRIBUTION

Conceptualization: Car-Emyr Suescum Coelho.
Data curation: Carluys Suescum Coelho, Carlysmar Suescum Coelho.
Formal analysis: Carlysmar Suescum Coelho.
Research: Car-Emyr Suescum Coelho, Carluys Suescum Coelho.
Methodology: Car-Emyr Suescum Coelho.
Project management: Carlysmar Suescum Coelho.
Supervision: Car-Emyr Suescum Coelho.
Visualization: Carlysmar Suescum Coelho.
Writing - original draft: Car-Emyr Suescum Coelho.
Writing - proofreading and editing: Carluys Suescum Coelho.