AHP-BASED FRAMEWORK FOR ASSESSING PROFESSIONAL TASK AUTOMATION RISK: METHODOLOGICAL DESIGN AND CROSS-SECTOR APPLICATION POTENTIAL

Sandra Patricia Barragán Moreno
Universidad de Bogotá Jorge Tadeo Lozano, Colombia
https://orcid.org/0000-0001-6503-4445
Bogotá, Colombia
Sandra.barragan@utadeo.edu.co

Gloria Patricia Calderón Carmona
Universidad de Bogotá Jorge Tadeo Lozano, Colombia
https://orcid.org/0000-0002-2176-702X
Bogotá, Colombia
gloria.calderon@utadeo.edu.co

Gabriel Budiño
Universidad de la República, Uruguay
https://orcid.org/0000-0002-1317-3379
gabriel.budino@fcea.edu.uy
Montevideo, Uruguay
gabriel.budino@fcea.edu.uy

ABSTRACT

The rapid advancement of artificial intelligence and automation is reshaping the nature of work; however, most occupational risk estimates remain aggregated, failing to account for the variability of automation susceptibility across specific tasks. This study introduces and validates a novel, transferable methodological framework and employs the Analytic Hierarchy Process (AHP) to assess automation risk at the task level, grounded in structured expert judgment. The framework decomposes an occupation into discrete tasks, evaluates these tasks against five systematically defined criteria—repetitiveness, cognitive complexity, human interaction, regulatory variability, and technological adaptability—and calculates priority vectors through pairwise comparisons. Internal consistency is verified, and inter-rater agreement is assessed using Kendall's coefficient. To evaluate the robustness of the results under expert variability, a Monte Carlo simulation-aided AHP (MC-AHP) approach was implemented, confirming the stability of task rankings. Applied to the accounting profession with an international panel of experts, the method identifies record keeping and report preparation as the most vulnerable tasks, while data analysis and management exhibit greater resilience. Beyond generating ranked vulnerability profiles, the approach captures diversity in expert perspectives, reflecting differences in regulatory and organizational contexts. The proposed framework offers decision-makers in diverse professional and geographical contexts a replicable, evidence-based tool for anticipating technological disruption and guiding workforce adaptation strategies in the era of digital transformation.

Keywords: Analytic Hierarchy Process; task automation; expert judgment; multi-criteria decision-making evaluation; automation risk

1. Introduction

The future of work is at the center of a global debate that has gained momentum following estimates suggesting that a significant percentage of current occupations could be affected by automation in the coming decades (Frey & Osborne, 2017; OIT, 2021). However, evidence shows that it is not entire occupations that are automated, but rather the specific tasks within them (Autor, Levy & Murnane, 2003; OIT, 2021). Previous research has demonstrated that routine and highly standardized tasks are more vulnerable to technological substitution, whereas those requiring professional judgment, human interaction, or contextual adaptation tend to be more resilient (Autor, Levy & Murnane, 2003; OIT, 2021). In Latin America, a region characterized by high levels of labor informality and persistent digital skills gaps (OIT, 2021), this distinction is crucial for guiding training policies, workforce reskilling, and business strategies.

Despite the growing attention to this issue, most automation risk models are constructed at the occupational level and seldom integrate expert judgment in a systematic way, which limits their ability to capture the diversity and complexity of tasks in specific contexts (ILO, 2021). This methodological gap calls for the development of flexible and replicable tools that allow automation risk to be assessed at the task level, incorporating both quantitative and qualitative criteria.

In this context, the Analytic Hierarchy Process (AHP) emerges as a multi-criteria decision-making technique capable of breaking down complex problems into hierarchies of criteria and alternatives, conducting pairwise comparisons, and synthesizing expert judgments into weighted priorities while verifying their consistency (Saaty, 1980). The AHP has been successfully applied in various fields, including the evaluation of occupational and educational risks, and its structure allows for the integration of key variables such as repetitiveness, cognitive complexity, human interaction, regulatory variability, and technological adaptability. However, given the growing number of publications and the diversity of applications of the AHP in this field, it is appropriate to conduct a systematic overview of the current literature to identify thematic trends, areas of concentration, and research gaps.

To this end, a mapping review was conducted, enabling the identification and visualization of both the structure and evolution of a research field through conceptual maps or bibliographic citation networks (Chambergo, Díaz & Benites, 2021). This approach provides a descriptive and visual overview of the available evidence and its bibliometric impact. The methodology was carried out in two stages: 1) The search and selection of articles in Scopus that met the criteria defined in the search equation; and 2) The synthesis of results based on the characteristics of the sample, keyword co-occurrence analysis, and clustering using VOSviewer. A cluster is defined as a group of nodes that are strongly related through their keyword co-occurrence (van Eck & Waltman, 2010). Given the scope of this work, no exclusion criteria were applied to the retrieved articles (Kraus, Breier & Lim, 2022).

The search returned a total of 17,188 documents using the following search equation:

ALL (("Analytic Hierarchy Process" OR "AHP") AND ("risk assessment" OR "risk evaluation" OR "hazard assessment"))
AND (LIMIT-TO (DOCTYPE , "ar"))
AND (LIMIT-TO (LANGUAGE , "English"))
AND (LIMIT-TO (SRCTYPE , "j"))
AND (LIMIT-TO (PUBSTAGE , "final"))
AND (LIMIT-TO (EXACTKEYWORD , "Risk Assessment") OR LIMIT-TO (EXACTKEYWORD , "Analytical Hierarchy Process")
OR LIMIT-TO (EXACTKEYWORD , "Analytic Hierarchy Process") OR LIMIT-TO (EXACTKEYWORD , "Hierarchical Systems")
OR LIMIT-TO (EXACTKEYWORD , "Risk Analysis") OR LIMIT-TO (EXACTKEYWORD , "Multicriteria Analysis")
OR LIMIT-TO (EXACTKEYWORD , "Risk Factor") OR LIMIT-TO (EXACTKEYWORD , "Risk Perception")
OR LIMIT-TO (EXACTKEYWORD , "Risks Assessments") OR LIMIT-TO (EXACTKEYWORD , "Analytic Hierarchy Process (ahp)")
OR LIMIT-TO (EXACTKEYWORD , "Risk") OR LIMIT-TO (EXACTKEYWORD , "Fuzzy Analytic Hierarchy Process")
OR LIMIT-TO (EXACTKEYWORD , "Risk Management"))

From this dataset, a keyword co-occurrence analysis was conducted using a threshold of 2. Figure 1 represents a keyword co-occurrence map, built from a set of articles indexed in an academic database. Each node represents a keyword, and its size reflects its frequency of occurrence. The connecting lines indicate co-occurrence relationships (i.e., how frequently keywords appear together). The colors identify thematic clusters (groups of terms that tend to appear together and thus represent a topic or subarea).

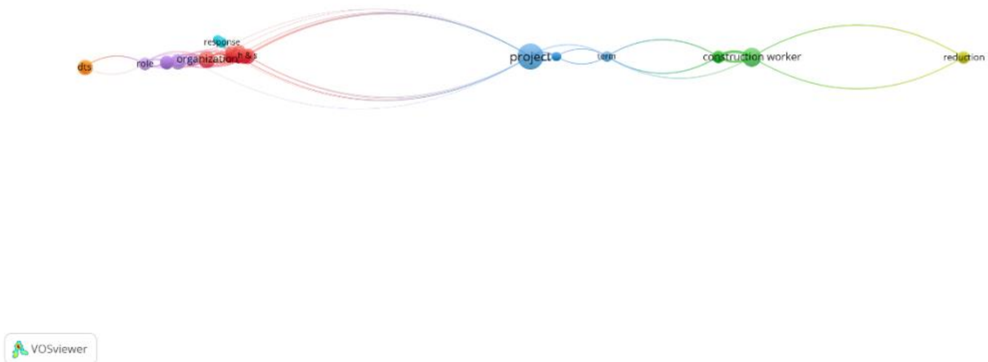

Figure 1 Keyword co-occurrence network based on Scopus-indexed literature on the Analytic Hierarchy Process (AHP) and risk-related studies (n = 17,188)
Source: Own elaboration using VOSviewer.

The analysis identified five main thematic clusters, reflecting the predominant areas in which AHP has been applied for risk-related studies:

Cluster 1: Organizational management and risk response (red)
This cluster includes terms such as organization, response, and role, and relates to studies that apply the AHP to support strategic decision-making in institutional settings, particularly in health management, disaster response, and organizational planning. The AHP is used to rank risk factors, allocate resources, and define response criteria under complex scenarios.

Cluster 2: Risk assessment in project environments (blue)
This group includes terms such as project, planning, and implementation, and is associated with research that applies the AHP in the analysis and management of risks throughout the life cycle of infrastructure, engineering, and information technology projects. The AHP supports the selection of alternatives, risk prioritization, and critical phase identification.

Cluster 3: Occupational risk and task-related hazards (green)
This cluster includes terms like construction worker, occupational risk, and safety, which correspond to studies focused on evaluating workplace risks in sectors such as construction, manufacturing, and mining. The AHP is applied to identify critical tasks, assess exposure conditions, and develop workplace safety strategies.

Cluster 4: Risk reduction and mitigation strategies (yellow)
Centered on the term reduction, this cluster includes studies that use the AHP to rank and select preventive and mitigation measures across various risk types (environmental, technological, financial). It highlights the use of the AHP in designing intervention plans, selecting control barriers, and evaluating response strategies to adverse events.

Cluster 5: Digital Transformation Strategies (DTS) (orange)
This emerging cluster, organized around the term DTS, includes studies that apply the AHP to assess risks related to digital transformation processes such as automation, AI integration, cybersecurity, and organizational change. The literature here reflects a growing interest in how the AHP can be used to support decision-making in digital risk environments and the reconfiguration of professional roles.

These clusters highlight both the thematic diversity and the methodological breadth of AHP applications in risk-related contexts. Furthermore, they reveal emerging lines of research (particularly around digital transformation) that point toward the method's continued evolution and relevance in addressing contemporary and future risk challenges.

Building on these findings, this study situates itself within the emerging research agenda identified in the literature, particularly the cluster focused on Digital Transformation Strategies (DTS). The prominence of this theme highlights the growing need for decision-making tools that address the risks posed by automation at a granular level. While the AHP has been widely applied across domains, its potential to evaluate automation risk at the task level in professional settings remains underexplored.

The aim of this study is to propose and validate an AHP-based methodological framework to assess the automation risk of professional tasks, illustrated through a case study in the accounting profession. This approach seeks to provide a more precise classification of the vulnerability of each task and to generate inputs for strategic decision-making in the context of accelerated digital transformation.

2. Reference framework

2.1 Automation, digitalization, and reconfiguration of work

Automation represents a structural transformation that redefines the division of labor between humans and machines. Its scope has expanded beyond industrial mechanization to include cognitive tasks—such as decision-making, data analysis, and customer service, enabled by advances in artificial intelligence, machine learning, and cyber-physical systems (Lorenz, Stephany & Kluge, 2023; Acemoglu & Restrepo, 2018). This transformation has given rise to a substantial body of interpretive literature, which encompasses diverse theoretical perspectives on the relationship between technology and employment. These perspectives can be systematized into five main approaches: substitution, complementarity, employment polarization, structural inequality, and task-

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

related risk (Frey & Osborne, 2017; Lorenz, Stephany & Kluge, 2023; Filippi, Bannò & Trento, 2023; Nedelkoska & Quintini, 2018; Bessen, 2018; Arntz, Gregory & Zierahn, 2016; Goos, Manning & Salomons, 2009; Autor, 2015; Autor & Dorn, 2013).

Table 1
Main theoretical approaches to understanding the impact of automation on employment

Theoretical approach	**Core description**
Substitution	Automation replaces human labor in tasks when it offers greater efficiency or cost-effectiveness.
Complementarity	Technology augments human capabilities, generating new functions that enhance productivity and value creation.
Employment polarization	Automation reduces middle-skill jobs while increasing demand for high-skill (low-risk) and low-skill (non-routine) roles.
Structural inequality	Technological change deepens disparities by economic status, education, and gender, especially in unequal markets.
Task-related risk	Automation risk is linked to the characteristics of individual tasks rather than entire occupations.

Source: Prepared by the authors based on Frey & Osborne, 2017; Lorenz, Stephany & Kluge, 2023; Nedelkoska & Quintini, 2018; Bessen, 2018; Arntz, Gregory & Zierahn, 2016; Goos, Manning & Salomons, 2009; Autor, 2015; Autor & Dorn, 2013.

One of the most widely recognized approaches is technological substitution, which posits that technical progress will replace human labor when machines become more efficient or cost-effective (Autor, 2015). Within this perspective, routine and codifiable tasks are the most susceptible to automation. Conversely, the complementarity approach argues that technology can augment human labor by enabling higher value-added activities and creating new functions that did not previously exist (Bessen, 2018).

The employment polarization perspective describes the simultaneous decline of middle-skill jobs and growth at both ends of the skill spectrum: high-skill occupations with low automation probability and low-skill jobs resistant to automation due to their interpersonal or physical nature (Goos, Manning & Salomons, 2009; Autor & Dorn, 2013). These trends are especially pronounced in unequal and fragmented labor markets, such as those in Latin America, where high levels of informality and limited access to technical training intensify automation's disruptive effects (Dicks, Künn-Nelen, Levels, & Montizaan, 2024; Alaimo, Chaves, & Soler, 2019).

2.2 Limitations of existing approaches to assessing automation risk

Automation risk analysis has evolved from occupation-based models to more granular, task-centered approaches. Frey and Osborne (2017) developed a pioneering model assigning automation probabilities to occupations using expert assessment on factors such as sensory perception, creativity, social intelligence, and object manipulation. OCDE

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

(2019) refined this approach by incorporating job skills survey data, producing more nuanced probability estimates.

Despite these advances, significant limitations remain. Many models rely on static assumptions that ignore evolving organizational structures and regulatory changes (Budiño & Asuaga, 2022). Structured integration of expert judgment is rare, and when present, it often lacks mechanisms to ensure consistency or to explicitly weight decision criteria. Furthermore, such models are seldom adaptable to the analysis of specific tasks within a single profession, and they frequently omit non-technical factors—such as social, regulatory, or organizational conditions—that can decisively shape automation outcomes (OIT, 2021; Filippi, Bannò & Trento, 2023; Schmidpeter & Winter-Ebmer, 2021). Table 2 summarizes these methodological limitations and their implications for policy and diagnostic tools.

Table 2
Key methodological limitations of common automation risk assessment models

Methodological limitation	**Implication**
Aggregate occupation analysis	Obscures intra-occupational variability in automation risk, reducing diagnostic granularity.
Lack of structured expert judgment	Reduces contextualization and validity of results for specific sectors.
Absence of consistency validation	Hinders reproducibility and comparability of findings across professions or countries.
Low adaptability to specific contexts	Limits the applicability of results for sector-specific policy design.
Exclusion of non-technical factors	Underestimates the influence of social, regulatory, and organizational determinants on technology adoption.

Source: Prepared by the authors based on Frey & Osborne, 2017; OIT, 2021; Filippi, Bannò & Trento, 2023; Budiño & Asuaga, 2022; Schmidpeter & Winter-Ebmer, 2021; Im & Schneider, 2022.

These constraints reduce the utility of current models for stakeholders in education, labor policy, and organizational planning, highlighting the need for methodologies that disaggregate occupations into tasks, evaluate automation risk across multiple dimensions, and incorporate expert judgment in a structured and replicable manner (Filippi, Bannò & Trento, 2023; Latham & Humberd, 2018).

2.3 AHP as a methodological framework for risk assessment

The AHP, developed by Saaty (2008), is a multi-criteria decision-making technique that breaks down complex problems into hierarchical levels. It facilitates pairwise comparisons among criteria and alternatives, synthesizing them into weighted priorities while also assessing the consistency of judgments. The AHP has been widely applied in various

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

domains to assess risk—including tourism and food supply chains, educational management, public policy evaluation, and occupational analysis—demonstrating its capacity to integrate both quantitative and qualitative criteria (Ali & Kaleem, 2025; Nguyen, Tuyen, Than & Tran, 2025; Barragán, González & Calderón, 2022).

In automation studies, the AHP has been used to assess the relative importance of tasks within an occupation, to prioritize risk factors such as task standardization, sensory perception, human interaction, and regulatory context, and to combine technical criteria with social perceptions about the future of work (Filippi, Bannò & Trento, 2023; Budiño & Asuaga, 2022). Its structured approach and capacity for replication make it valuable for comparative analyses across professions and sectors.

2.4 Justification for using the AHP for automation risk assessment

Applying the AHP to automation risk assessment addresses critical limitations in previous models. First, it enables analysis at the task level rather than the occupation level, providing more precise and actionable insights. Second, its flexible structure allows adaptation to varied professional and regulatory contexts. Third, it facilitates participatory validation by incorporating expert knowledge in a systematic way, ensuring multiple perspectives are integrated and evaluated for consistency (Maskavizan, Poco & Calzolari, 2023; Kendall, 1938).

By linking disaggregated tasks to criteria such as routinization, cognitive complexity, human interaction, and technological adaptability, the AHP produces a hierarchical model that accurately reflects an occupation's susceptibility to automation. This methodological approach offers decision-makers and educational stakeholders a robust tool for curriculum design, professional retraining, and evidence-based policymaking in the context of accelerated digital transformation.

3. Methodology

3.1 Study design

This study is methodological in nature and proposes a systematic, transparent, and replicable framework for assessing the automation risk of professional tasks using the AHP. The framework adapts the AHP to structure expert judgments regarding the susceptibility of tasks to automation through hierarchical criteria, enabling their quantitative synthesis.

The approach is quantitative, as it transforms qualitative expert assessments into pairwise comparison matrices processed through mathematical algorithms. This structure supports the calculation of the relative weighting of specific tasks, the evaluation of consistency in expert opinions, and the generation of composite automation risk scores.

Although primarily methodological, the framework is validated through its application to the accounting profession, serving as an exemplary case to demonstrate its implementation

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

and effectiveness. This validation aims to illustrate the approach rather than produce generalizable empirical conclusions.

3.2 Application of the AHP to automation risk assessment

A task-based approach was adopted, decomposing each occupation into discrete tasks potentially susceptible to automation (Lorenz, Stephany & Kluge, 2023; Budiño & Asuaga, 2022). This required explicit weighting of task types and the factors influencing their likelihood of automation. The AHP integrates multiple criteria—such as required skills and task characteristics—into a single decision-making model, assigning weights according to their contribution to automation risk.

The AHP's methodological flexibility allows for the integration of qualitative and quantitative factors in a unified framework (Anderson, Sweeny, & Williams, 2016) and offers mechanisms to evaluate the consistency of expert judgments (Barragán, González & Calderón, 2022; Davoodi, 2009).

3.3 Definition of evaluation criteria and sub-criteria

Key factors influencing automation risk were identified from the literature on technological change and work transformation (Chambergo, Díaz & Benites, 2021; Barragán & Guzmán, 2025; Grant & Booth, 2009). Each criterion captures a relevant dimension of automation risk, as shown in Table 3.

Table 3
Hierarchy of automation risk assessment criteria of professional tasks

Criterion	Sub-criteria	Theoretical justification
Repetitiveness of tasks	Routine nature of tasks, process standardization	Highly repetitive and structured tasks are easier to automate.
Requires cognitive complexity	Creativity, complex problem solving	Tasks that require professional judgment and non-standard problem-solving are less susceptible to automation.
Human interaction and empathy	Social/emotional intelligence, interpersonal communication	Human interaction, empathy, and negotiation are difficult for algorithms to replicate.
Physical-manual dexterity and sensory perception	Sensory perception, fine motor coordination	Some manual tasks require adaptive skills that are difficult to automate.

Source: Prepared by the authors based on Frey & Osborne, 2017; Budiño & Asuaga, 2022; Latham & Humberd, 2018; Włoch, K. Śledziewska & Rożynek, 2005; Bannò, Filippi & Trento, 2021.

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

3.4 Hierarchical model design

The model has three levels: 1) General objective – Assess the automation risk of tasks in each profession; 2) Criteria (risk factors) – Derived from the literature and empirical evidence (see Table 3); and 3) Decision alternatives (tasks) – Specific tasks evaluated for automation susceptibility.

This structure enables flexible application across occupations, supporting multi-criteria evaluation grounded in expert judgment (Anderson, Sweeny, & Williams, 2016).

3.5 Pairwise comparison process and weight calculation

Once criteria are defined, experts conduct pairwise comparisons using Saaty's scale (Table 4), producing matrices at each hierarchy level.

Table 4
Pairwise comparison scale for the AHP preferences

Verbal judgment	Numerical value
Equally preferred	1
From equally to moderately preferred	2
Moderately more preferred	3
From moderately to strongly preferred	4
Strongly more preferred	5
From strongly to very strongly preferred	6
Very strongly more preferred	7
From very strongly to extremely preferred	8
Extremely more preferred	9

Source: Prepared by the authors based on Saaty, 1980; Saaty, 2008.

3.6 Priority estimation algorithm with AHP

To operationalize the AHP for automation risk assessment, the calculation of priorities follows the sequence established in classical literature (Anderson, Sweeny & Williams, 2016; Munier, 2011). The process is designed to progressively translate expert judgments into quantitative measures that can be systematically compared and synthesized.

Step 1 – Identification of criteria and risk factors: The first stage defines the key risk criteria/factors at Level 1 of the hierarchy. These are derived from the literature review and represent the main dimensions influencing the susceptibility of professional tasks to automation.

Step 2 – Construction of pairwise comparison matrices: A square matrix A is built, where a_ij denotes the expert's judgment of the relative importance of element i compared to element j. The value $a_{ij} = 1$ indicates equal importance. Two levels of matrices are constructed:

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

- One comparing the main risk criteria/factors.
- Multiple matrices comparing the decision alternatives/tasks under each criterion.

Step 3 – Estimation of relative weights: The normalized principal eigenvector of each matrix is calculated to determine the weight of each criterion or alternative. This step is repeated for every matrix, ensuring that each set of comparisons is internally consistent and that the weights reflect the experts' collective understanding of the problem structure.

Step 4 – Weighting of alternatives: For each decision alternative (task), the local priority values obtained in Step 3 are multiplied by the weight of the corresponding criterion. This integration ensures that both the intrinsic importance of the criterion and the performance of the alternative under that criterion are reflected.

Step 5 – Aggregation into composite scores: The weighted values are summed across all criteria, yielding a composite score for each task. This score serves as a single, interpretable indicator of automation risk, facilitating comparison and prioritization.

Step 6 – Consistency verification: The internal consistency of expert judgments is evaluated by computing the Consistency Ratio (CR). This step ensures that the judgments are logically coherent and that the resulting weights are reliable.

It is important to note that Steps 2 to 5 are executed independently for each referee, producing an individual priority matrix that reflects each expert's perspective. These individual results are subsequently aggregated (via the geometric mean) into a consolidated priority ranking. This procedure not only preserves the diversity of expert input but also ensures methodological rigor in line with multi-criteria decision-making best practices.

3.7 Evaluation of the consistency of judgments

Ensuring the logical coherence of expert judgments is an essential requirement for the reliability of the AHP results. Once the priority vectors have been calculated, the Consistency Ratio (CR) is used to evaluate whether the pairwise comparisons provided by each referee are consistent with the principles of transitivity and proportionality in decision-making.

The verification process follows the algorithm below:

Step 1 – Weighted sum vector calculation: Multiply the pairwise comparison matrix by the priority vector to obtain the weighted sum vector.
Step 2 – Ratio calculation: Divide each element of the weighted sum vector by its corresponding priority value.
Step 3 –Maximum eigenvalue (λ max) determination: Calculate the mean of the values obtained in Step 2. This mean represents the maximum eigenvalue of the matrix.
Step 4 – Consistency Index (CI) computation: Compute the Consistency Index using the following equation, where n is the number of elements compared:

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

$$CI = \frac{\lambda_{max-n}}{n-1} \quad (1)$$

Step 5 – Consistency Ratio (CR) computation: Obtain the CR by dividing the CI by the Random Index (RI) corresponding to n, as shown in the equation:

$$CR = \frac{CI}{RI} \quad (2)$$

The Random Index (RI) values used in this study are presented in Table 5. These values are derived from randomly generated reciprocal matrices of size n (Mendoza, 2013).

Table 5
Consistency index of a randomly generated matrix

$\boldsymbol{n}$	3	4	5	6	7	8	9
$\boldsymbol{RI}$	0.6	1	1.1	1.2	1.3	1.41	1.45

Source: Mendoza (2013, p. 37).

Interpretation: A $CR \leq 0.10$ is generally considered acceptable, indicating that the judgments have a satisfactory level of internal coherence. If the CR exceeds this threshold, Mendoza (2013) recommends reviewing the expert's comparisons for possible inconsistencies. In some cases, and if justified by the decision-making context, a more flexible threshold may be applied. For alternative methods of assessing the consistency of expert responses, see Barragán, González and Calderón (2022) and Alonso and Lamata (2006).

3.8 Agreement among referees

Beyond verifying the internal consistency of each referee's judgments, it is essential to determine the degree of consensus within the panel of experts. Discrepancies in the assessment of the relative importance of criteria/risk factors or decision alternatives/tasks can influence the robustness of the aggregated results.

To quantify the level of agreement, this study employed the Kendall Coefficient of Concordance (W), a non-parametric statistic that measures the ordinal association among multiple raters (Maskavizan, Poco & Calzolari, 2023; Kendall, 1938; Arízaga Piedra, Monge Loja & Muñoz Pauta, 2023). The coefficient ranges from 0 (no agreement, equivalent to random ranking) to 1 (perfect agreement).

The algorithm for calculating Kendall's W is as follows:

Step 1 – Matrix construction: Create an $n \times m$ matrix, where n represents the number of decision alternatives/tasks and m the number of referees. Each cell P_{ij} contains the rank assigned by referee j to alternative/task i.

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

Step 2 – Sum of ranks: For each alternative/task, calculate the sum of its ranks:

$$H_i = \sum_{j=1}^{m} P_{ij} \tag{3}$$

Step 3 – Mean of total ranks: Compute the mean of the rank sums:

$$\bar{H} = \frac{1}{n}\sum_{i=1}^{n} H_i \tag{4}$$

Step 4 – Sum of squared deviations (S): Determine the dispersion of the rank sums around the mean:

$$S = \sum_{i=1}^{n} (H_i - \bar{H})^2 \tag{5}$$

Step 5 – Kendall's Coefficient of Concordance (W): Calculate W using:

$$W = \frac{12S}{m^2(n^3 - n)} \tag{6}$$

Interpretation:
- $W = 1$ Perfect agreement among referees
- $W = 0$ No agreement (random ordering)
- $W < 0.3$ Slight agreement
- $0.3 \leq W < 0.7$ Moderate agreement
- $W \geq 0.7$ High agreement

The computation of W provides a quantitative measure of consensus, supporting the validation of aggregated rankings. High levels of agreement strengthen the reliability of the results, while lower values may indicate the need to revisit criteria definitions, clarify task descriptions, or provide additional guidance to referees before re-assessment.

3.9 Final prioritization

Once the individual priority vectors for each referee have been obtained, both for the criteria/risk factors and for the decision alternatives/tasks, the next stage is to integrate these results into a consolidated ranking that reflects the collective judgment of the expert panel.

To achieve this, the geometric mean was applied to the corresponding entries of the individual priority vectors. This approach, widely recommended in the AHP literature for aggregating individual judgments, ensures that the influence of each referee is proportionally represented and that the aggregated priorities preserve the multiplicative nature of pairwise comparisons.

The consolidation procedure follows these steps:

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

Step 1 – Aggregation of criteria priorities: For each criterion/risk factor, compute the geometric mean of the relative weights assigned by all referees. This yields a global criteria priority vector.
Step 2 – Aggregation of task priorities within each criterion: For each decision alternative/task under each criterion, calculate the geometric mean of the relative weights assigned by all referees, generating global task priority vectors for each criterion.
Step 3 – Weighted aggregation across criteria: Multiply the priority of each task (within a given criterion) by the global weight of its corresponding criterion.
Step 4 – Final composite score calculation: Sum the weighted priorities of each task across all criteria to obtain a composite automation risk score for each decision alternative/task.
Step 5 – Ranking: Order the tasks in descending order of their composite scores to identify those with the highest and lowest susceptibility to automation.

This procedure ensures that the final prioritization: 1) Incorporates the judgment of all referees in a balanced manner; 2) Preserves the methodological integrity of the AHP process; 3) Produces a transparent and reproducible ranking of professional tasks; 4) Provides a decision-making tool that can be directly used to identify critical tasks for job redesign, workforce training, or automation mitigation strategies. In the context of the case study on the accounting profession, these results also offer insights into which activities require urgent strategic attention, and which are less vulnerable to technological substitution.

3.10 Implementation in a case study: accounting profession

To demonstrate the applicability and robustness of the proposed AHP-based framework, the methodology was operationalized in a case study focusing on the accounting profession, a field identified in the literature as highly susceptible to automation (Frey & Osborne, 2017; Lorenz, Stephany & Kluge, 2023). This case provides a relevant and challenging testing ground for the method, as it includes both routine, rule-based activities (such as transaction recording, reconciliations, and standardized reporting) that exhibit a high probability of automation, and analytical, judgment-based, or interpersonal tasks that remain less susceptible to technological replacement (Budiño & Asuaga, 2022).

The selection of this profession is further supported by its accelerated adoption of Robotic Process Automation (RPA) and artificial intelligence (AI) tools, which have already redefined operational workflows in many accounting contexts (Ogosi, Lira, Guadalupe, Lira & Cacsire, 2023; Aguirre & Rodríguez, 2017). As such, it offers a realistic environment to test the framework in a sector where technological disruption is actively reshaping task structures.

Following the methodological sequence described in previous paragraphs, the implementation phase comprised:

1. Application of the hierarchical criteria defined for automation risk assessment, adapted to the specific tasks and competencies of accounting professionals.

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

2. Collection of expert judgments from an international panel, systematically structured through pairwise comparisons in accordance with the AHP methodology.
3. Evaluation of internal consistency and inter-rater agreement, using the algorithms described in the consistency and concordance sections, to ensure methodological reliability and transparency.

This structured approach aligns with best practices in methodological validation research (Landriscina, 2013) integrating theoretical model development, expert-based data collection, and robust statistical verification. The results not only confirm the framework's practical utility for assessing automation risk but also generate actionable insights for workforce planning, occupational redesign, and targeted training policies in accounting and potentially in other professional domains.

3.11 Sensitivity analysis using Monte Carlo simulation

To evaluate the robustness of the rankings under expert judgment variability, a sensitivity analysis was conducted using a Monte Carlo simulation–aided Analytic Hierarchy Process (MC-AHP). Following the methodology established by Jing, Chen, Zhang, Li, and Zheng (2013), applied by Thomas, Rane, and Harris (2024) in nuclear risk and security studies, and extended to multi-stakeholder decision-making by Mantogiannis and Katsigiannis (2020), the following four steps were undertaken:

Step 1 – Triangular estimation: For each non-diagonal element of the pairwise comparison matrices, a set of three values—minimum (min), most likely (modal), and maximum (max)—was derived from the judgments provided by the three experts. The minimum and maximum correspond to the lowest and highest observed values, respectively, while the modal value corresponds to their arithmetic mean. This structure captures the range of expert input and enables probabilistic modeling of uncertainty.

Step 2 – Beta-PERT distribution modeling: Using the estimated triplets, beta-PERT distributions were constructed to model the uncertainty of each judgment. The beta-PERT distribution is well suited for expert-based estimation because it allows for skewed distributions and emphasizes the most likely value while accounting for variability. The following equations (adapted from Arntz, Gregory & Zierahn, 2016) were used to parameterize the distribution:

$$mean = \frac{min+4modal+max}{3} \tag{7}$$

$$stdev = \frac{max-min}{3} \tag{8}$$

$$\alpha = \left(\frac{mean-min}{max-min}\right)\left(\frac{(mean-min)(max-min)}{stdev^2} - 1\right) \tag{9}$$

$$\beta = \alpha\left(\frac{max-mean}{mean-min}\right) \tag{10}$$

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

These shape parameters were then used to simulate the non-diagonal values in the pairwise matrices through random sampling from the beta distribution.

Step 3 – Simulation setup and score aggregation: A Python-based simulation model was developed to generate 1,000 synthetic realizations of each pairwise comparison. For each iteration, the model sampled values from the beta-PERT distributions, constructed the complete pairwise matrix (enforcing reciprocity), and calculated the priority vector using the eigenvector method.

The global score $A_k$for each alternative $k$was then calculated by aggregating the local scores b_{kj}(task scores under each criterion j) weighted by the normalized importance $U_j$of the five criteria, following the approach used by Thomas, Rane, & Harris (2024):

$$A_k = \sum_{j=1}^{5}\left(b_{kj} \cdot U_j\right) \tag{11}$$

Step 4 – Result synthesis and visualization: The simulation produced a distribution of 1,000 scores per task. From these distributions, statistical descriptors such as mean, standard deviation, and 95% confidence intervals were computed. Additionally, boxplots were generated to visually assess the dispersion of scores and the relative stability of task rankings under expert judgment variability.

This simulation-based sensitivity analysis offers a robust and transparent means of evaluating the stability of AHP-based decisions. It not only quantifies the potential impact of uncertainty in expert input but also enhances the methodological rigor and replicability of AHP applications in this risk assessment context.

4. Results

This section presents the findings from the implementation of the AHP-based methodology designed to assess the automation risk of professional tasks, validated in the accounting profession. As detailed in the methodological section, the hierarchical structure of the model was applied with input from an international panel of experts. These experts provided pairwise comparative judgments at two levels: 1) Among the criteria/risk factors, and 2) Among the decision alternatives/tasks within each criterion.

Individual judgments were consolidated using the geometric mean, producing a consensus-based priority ranking in line with AHP best practices (Anderson, Sweeny & Williams, 2016; Munier, 2011).

4.1 Characterization of the panel of experts

Three international experts with extensive experience in technological automation and the accounting field were selected as referees. Selection criteria included:

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

- Thematic specialization – All experts had in-depth knowledge of the current capabilities and limitations of technologies applied to accounting and work organization, enabling nuanced understanding of professional tasks and competencies.
- Academic or professional affiliation – Each expert was affiliated, at least partially, with a higher education institution.
- Geographic diversity – Experts came from Colombia, Argentina, and Uruguay, allowing incorporation of diverse perspectives shaped by economic and regulatory contexts (Budiño & Asuaga, 2022).

This diversity mitigated potential local bias, given that perceptions and responses to automation vary significantly across countries (Buzzelli, 2023).

4.2 Definition and prioritization of evaluation criteria

Based on the framework presented in Table 3 and the professional expertise of two accounting specialists from Uruguay and Colombia, the criteria/risk factors used to assess the automation risk in the accounting profession were defined as follows:

1. Repetitiveness: Tasks with fixed patterns and low variability—requiring minimal judgment—are highly automatable through technologies like RPA (Frey & Osborne, 2017).
2. Complexity: Tasks involving critical thinking, judgment, and decisions in ambiguous contexts are less automatable due to their reliance on tacit knowledge (Acemoglu & Restrepo, 2018).
3. Reliance on social skills or human interaction: Functions requiring social skills such as empathy or negotiation resist automation, as emotional and interpersonal nuances are hard to replicate (Autor, 2015).
4. Variability in the regulatory context: Tasks that demand continuous adaptation to evolving legal frameworks are less standardizable and thus less automatable on (Włoch, Śledziewska & Rożynek, 2005).
5. Adaptability to emerging technology: Tasks easily integrated with emerging technologies like AI or big data are more prone to automation due to their high compatibility (Bessen, 2018).

4.3 Definition of decision alternatives

Following Latham and Humberd (2018) the accounting tasks were grouped into four decision alternatives:

1. Record Keeping – Manual, routine activities such as data entry and reconciliations
2. Reporting – Routine cognitive tasks like adjustments and data verification
3. Management – Non-routine cognitive tasks involving analysis and design of processes
4. Data Analysis – High-complexity, non-repetitive tasks such as consulting and decision-making

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

4.4 Hierarchical model structure

Figure 2 illustrates the AHP-based hierarchical structure for evaluating automation risk in the accounting profession, linking the overall objective, criteria/risk factors, and decision alternatives/tasks.

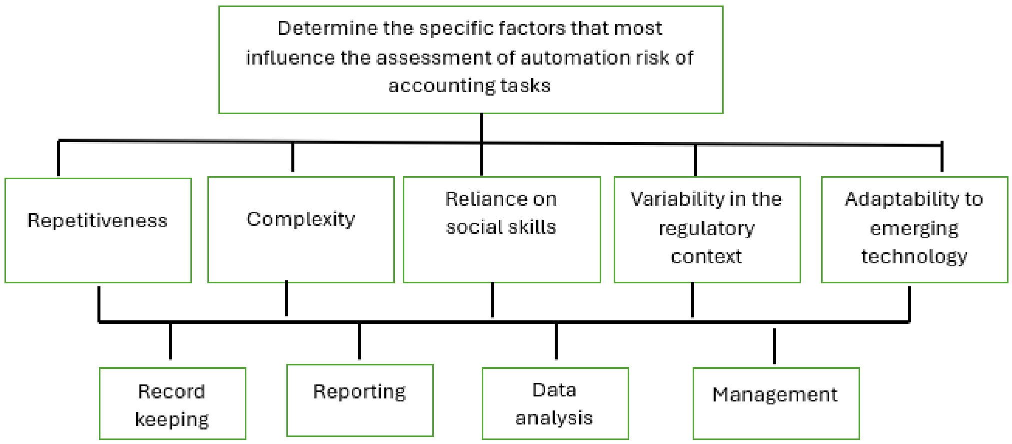

Figure 2 Hierarchical structure of the AHP-based model for assessing automation risk in the accounting profession

4.5 Priority vectors for criteria and alternatives

The three referees provided judgments on both criteria and tasks, producing the priority vectors v_{cRi}, $i = 1,2,3$. for criteria and v_{cR1} for tasks (Table 6). For example, Referee 1 assigned the highest weight (0.58) to Repetitiveness, while Referee 2 assigned greater emphasis (0.25) to Complexity.

$$v_{cR1} = \begin{pmatrix} 0.58 \\ 0.14 \\ 0.15 \\ 0.06 \\ 0.04 \end{pmatrix}; v_{cR2} = \begin{pmatrix} 0.32 \\ 0.25 \\ 0.14 \\ 0.10 \\ 0.17 \end{pmatrix}; v_{cR3} = \begin{pmatrix} 0.48 \\ 0.21 \\ 0.12 \\ 0.10 \\ 0.06 \end{pmatrix}$$

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

Table 6
Priority vectors assigned by each referee to accounting tasks under each risk criterion

Criterion	Referee 1	Referee 2	Referee 3
Repetitiveness	$v_{T1R1} = \begin{pmatrix} 0.67 \\ 0.17 \\ 0.06 \\ 0.08 \end{pmatrix}$	$v_{T1R2} = \begin{pmatrix} 0.60 \\ 0.24 \\ 0.11 \\ 0.03 \end{pmatrix}$	$v_{T1R3} = \begin{pmatrix} 0.63 \\ 0.22 \\ 0.09 \\ 0.04 \end{pmatrix}$
Complexity	$v_{T2R1} = \begin{pmatrix} 0.36 \\ 0.36 \\ 0.13 \\ 0.12 \end{pmatrix}$	$v_{T2R2} = \begin{pmatrix} 0.54 \\ 0.27 \\ 0.13 \\ 0.04 \end{pmatrix}$	$v_{T2R3} = \begin{pmatrix} 0.59 \\ 0.22 \\ 0.14 \\ 0.04 \end{pmatrix}$
Reliance on social skills or human interaction	$v_{T3R1} = \begin{pmatrix} 0.37 \\ 0.29 \\ 0.23 \\ 0.09 \end{pmatrix}$	$v_{T3R2} = \begin{pmatrix} 0.59 \\ 0.24 \\ 0.12 \\ 0.03 \end{pmatrix}$	$v_{T3R3} = \begin{pmatrix} 0.59 \\ 0.23 \\ 0.12 \\ 0.04 \end{pmatrix}$
Variability in the regulatory context	$v_{T4R1} = \begin{pmatrix} 0.65 \\ 0.21 \\ 0.06 \\ 0.06 \end{pmatrix}$	$v_{T4R2} = \begin{pmatrix} 0.36 \\ 0.36 \\ 0.16 \\ 0.09 \end{pmatrix}$	$v_{T4R3} = \begin{pmatrix} 0.34 \\ 0.34 \\ 0.21 \\ 0.09 \end{pmatrix}$
Adaptability to emerging technology	$v_{T5R1} = \begin{pmatrix} 0.46 \\ 0.27 \\ 0.14 \\ 0.11 \end{pmatrix}$	$v_{T5R2} = \begin{pmatrix} 0.56 \\ 0.26 \\ 0.12 \\ 0.03 \end{pmatrix}$	$v_{T5R3} = \begin{pmatrix} 0.56 \\ 0.25 \\ 0.13 \\ 0.04 \end{pmatrix}$

Note: Relative priorities (sum = 1.0) assigned by each referee; higher values indicate greater perceived automation risk within the criterion.

4.6 Consistency of judgments

Table 7 shows the Consistency Index (CI) for each referee at both criteria and task levels. Values above the 0.10 threshold indicate potential inconsistency (Mendoza, 2013). Referee 1 maintained high consistency (global CI = 0.07), although CI rose to 0.13 under Adaptability to Emerging Technology.

Table 7
Consistency Index (CI) of referees' judgments

CI	Referee 1	Referee 2	Referee 3
Criteria	0.07	0.19	0.08
Repetitiveness	0.06	0.58	0.19
Complexity	0.09	0.45	0.17
Reliance on social skills or human interaction	0.14	0.50	0.28
Variability in the regulatory context	0.03	0.14	0.02
Adaptability to emerging technology	0.13	0.35	0.27

4.7 Aggregation and concordance

Using the priority vectors for each criterion and task, weighted matrix multiplication yielded the classification vectors $P_{Ri} = (v_{t1Ri}|v_{t2Ri}|v_{t3Ri}|v_{t4Ri}|v_{t5Ri})v_{cRi}$

For example, Referee 1 ranked tasks as: 1) Record Keeping (0.57); 2) Reporting (0.22); 3) Data Analysis (0.10); and 4) Management (0.09).

$$P_{R1} = \begin{pmatrix} 0.57 \\ 0.22 \\ 0.10 \\ 0.09 \end{pmatrix};\ P_{R2} = \begin{pmatrix} 0.55 \\ 0.27 \\ 0.12 \\ 0.04 \end{pmatrix};\ P_{R3} = \begin{pmatrix} 0.58 \\ 0.24 \\ 0.12 \\ 0.04 \end{pmatrix}$$

To evaluate agreement among referees, Kendall's coefficient of concordance (W) was calculated using Equations (3) to (6) (Table 8). The resulting W = 0.031 indicates low agreement, highlighting differences in experts' prioritization—possibly reflecting national and sectoral perspectives (Buzzelli, 2023).

Table 8
Calculation of Kendall's coefficient of concordance

	Prioritization matrix			Sum of priority vectors	Sum of squared deviations	
Tasks	P_{R1}	P_{R2}	P_{R3}	H_i	$H_i - \overline{H}$	$(H_i - \overline{H})^2$
Record Keeping	0.57	0.55	0.58	1.71	0.96	0.93
Reporting	0.22	0.27	0.24	0.73	-0.01	0.00
Data Analysis	0.10	0.12	0.12	0.35	-0.39	0.15
Management	0.09	0.04	0.04	0.18	-0.56	0.31

Note: The mean of the sums was $\overline{H}$=0.75, and the sum of squared deviations was $S = 1.4$.

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

4.8 Consensus-based ranking

The final consensus vector was computed via the geometric mean of referees' scores, producing the ranking: 1) Record Keeping (0.56); 2) Reporting (0.24); 3) Data Analysis (0.11); and 4) Management (0.01). Tasks with higher complexity, contextual variability, and reliance on interpersonal skills (e.g., Data Analysis, Management) showed lower automation risk, consistent with previous findings (Acemoglu & Restrepo, 2018; Autor, 2015).

$$P = \begin{pmatrix} 0.56 \\ 0.24 \\ 0.11 \\ 0.01 \end{pmatrix}$$

The resulting vector reflects the consensus-based ranking of tasks according to the relative automation risk, as determined by the multi-criteria decision-making approach proposed in this study. The findings indicate that, among the tasks evaluated, Record Keeping is the most susceptible to automation (0.56), followed by Reporting (0.24). In contrast, tasks such as Data Analysis (0.11) and Management (0.01) are considered to have substantially lower risk of automation. This is attributed to their higher cognitive complexity, contextual variability, and reliance on interpersonal skills. Overall, the classification highlights the heterogeneous nature of accounting tasks in the face of technological advances and supports the relevance and applicability of the proposed multi-criteria decision-making approach for assessing automation risks in a rigorous and context-sensitive manner.

Interpretation: The results confirm that the proposed AHP-based methodology is effective for identifying tasks with greater susceptibility to automation. The heterogeneous ranking across task types reflects the multidimensionality of automation risk in accounting, supporting the relevance, flexibility, and transferability of this approach to other professional contexts.

4.9 Sensitivity analysis results

To assess the robustness of the task prioritization results under expert judgment variability, a Monte Carlo simulation–aided AHP (MC-AHP) was conducted. The simulation followed the approach proposed by Jing, Chen, Zhang, Li, and Zheng (2013), using beta-PERT distributions for each non-diagonal element of the pairwise comparison matrices based on the minimum, modal (mean), and maximum values obtained from three experts. A total of 1,000 iterations were performed, generating synthetic realizations of the judgment matrices and computing the global priority score for each task per iteration using the normalized weights of the five criteria.

Table 9 summarizes the statistical descriptors obtained from the simulation. The results indicate that Record Keeping consistently received the highest automation risk score (mean = 0.5646; SD = 0.0123), followed by Reporting (mean = 0.2486; SD = 0.0081), Data Analysis (mean = 0.1213; SD = 0.0051), and Management (mean = 0.0654; SD = 0.0052).

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

The observed 95% confidence intervals were narrow and did not overlap across tasks, which supports the consistency of the ranking despite variability in individual judgments.

Table 9
Statistical summary of global priority scores from MC-AHP (n = 1,000)

Task	Mean	Std. Dev.	Min	2.5%	Median	97.5%	Max
Record Keeping	0.5646	0.0123	0.5253	0.5401	0.5648	0.5883	0.6000
Reporting	0.2486	0.0081	0.2218	0.2325	0.2485	0.2642	0.2737
Data Analysis	0.1213	0.0051	0.1046	0.1114	0.1213	0.1313	0.1368
Management	0.0654	0.0052	0.0507	0.0559	0.0654	0.0758	0.0798

Figure 3 presents the boxplots of the global scores obtained for each task across all simulations. The visual distribution of scores confirms the low dispersion and reinforces the stability of the ranking. Although "Record Keeping" displays the widest range, it maintains the highest risk level throughout all iterations. This pattern is consistent with the deterministic results, confirming that the prioritization is robust under uncertainty.

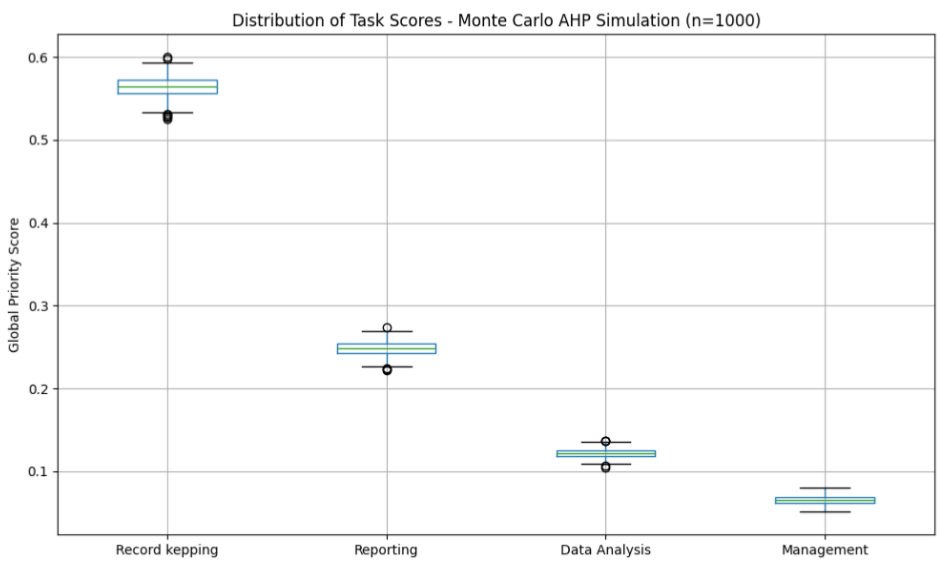

Figure 3 Distribution of global priority scores for professional tasks obtained from Monte Carlo simulation–aided AHP (n = 1,000 iterations)

This simulation-based sensitivity analysis addresses the methodological limitations often attributed to traditional AHP models, particularly the absence of statistical validation and judgment uncertainty (Thomas, Rane & Harris, 2024). By explicitly modeling variability in expert input, the MC-AHP approach enhances the reliability and transparency of task-level risk assessments.

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

5. Discussion

5.1 Discussion of the results

This study's primary contribution lies in demonstrating the methodological value of the Analytic Hierarchy Process (AHP) for assessing automation risk in professional tasks through expert judgment. Rather than focusing on the substantive results for a specific profession, the emphasis is on how the proposed procedure enables the decomposition of complex decision problems into structured, conceptually clear levels, facilitating systematic comparisons between criteria and alternatives—even in highly complex contexts (Saaty, 2008; Munier, 2011). The design of clear and guided evaluation instruments was critical to ensure that experts fully understood the task, an essential factor when incorporating specialized knowledge into multi-criteria decision-making processes (Alonso & Lamata, 2006).

The case study in the accounting profession serves as a validation scenario for the methodology. The Kendall coefficient of concordance indicated a slight level of agreement among referees (Maskavizan, Poco & Calzolari, 2023; Kendall, 1938), consistent with expectations for expert-based multi-criteria assessments. While individual perceptions varied, shared ranking patterns emerged, enabling the identification of tasks with the highest automation risk—namely Record Keeping and Reporting. These tasks are repetitive and predictable, characteristics that previous studies have consistently linked to technological substitution (Frey & Osborne, 2017; Arntz, Gregory & Zierahn, 2016; Budiño & Asuaga, 2022). This outcome confirms the ability of the AHP-based framework to capture widely recognized automation risk patterns, while also allowing the identification of context-specific nuances.

The diversity of the expert panel—though limited to South American academics—illustrates how the method integrates both global and local perspectives. Judgments reflected broader trends in automation as well as the specific regulatory and operational realities of accounting practice in the participating countries. This alignment with the recommendations of Alaimo, Chaves, and Soler (2019) and Ripani, Soler, Kugler, Kugler, and Rodrigo (2020) strengthens the case for adapting analytical frameworks to diverse social, labor, and technological contexts.

The methodological design also benefited from aggregating expert judgments using the geometric mean, which preserved individual independence and reduced potential group bias (Davoodi, 2009; Alonso & Lamata, 2006). This not only increased the robustness of the rankings but also yielded a transparent and replicable procedure that can be transferred to other fields facing automation challenges.

5.2 Theoretical implications

The use of the AHP as a methodological framework reinforces the argument that structured decision-making tools are critical in contexts characterized by complexity and uncertainty. In this study, the AHP proved to be a powerful tool for capturing both generalizable trends and profession-specific nuances in automation risk, particularly in domains where

ISSN 1936-6744
https://doi.org/10.13033/ijahp.v17i3.1370

technical feasibility interacts with contextual variables such as legal mandates or task complexity. The findings contribute to the growing literature on the intersection between digital transformation and occupational change, demonstrating how expert knowledge can be systematically translated into hierarchical criteria. Moreover, the study illustrates that the incorporation of diverse expert perspectives, even within a regional academic context, allows for valid theoretical insights into how automation risk is socially constructed and contextually grounded.

5.3 Practical implications

The proposed AHP-based framework offers practical value to policymakers, educators, and organizational leaders who must anticipate the effects of automation on professional roles. For instance, identifying Record Keeping and Reporting as the most automatable accounting tasks provides a foundation for redesigning training programs and job roles to prioritize higher-order skills that are less susceptible to automation. Additionally, organizations can adapt the methodological design to other professions or sectors where similar pressures from digitalization are present. The use of transparent and replicable processes enhances stakeholder trust and supports evidence-based workforce planning. The insights on local regulatory influence also encourage a nuanced approach to automation policies, one that accounts for institutional and legal variability across countries or jurisdictions.

5.4 Limitations and future research

The methodology's public dissemination enables continuous refinement, adaptation to new professional contexts, and incorporation of evolving technological trends. Nonetheless, certain limitations merit acknowledgment. First, the validation panel, although professionally relevant, was relatively small and exclusively academic. This composition may limit generalizability, suggesting that future studies expand the panel to include a more diverse and international pool of experts or adopt Delphi rounds to reinforce independence and robustness of judgments. Second, the pairwise comparison scale may introduce anchoring bias if referees avoid extreme values on Saaty's scale. Although the research team provided clear instructions to mitigate this risk, complementary calibration exercises could further enhance judgment precision. Finally, the approach measures technical feasibility of automation but does not explicitly account for economic, organizational, or regulatory constraints that may impede implementation. Tasks deemed technically automatable may remain in practice due to high implementation costs, protective regulations—such as the legal role of public accountants as notaries in Colombia—or institutional resistance to change.

To test the stability of the results under conditions of expert judgment variability, a Monte Carlo simulation–aided AHP (MC-AHP) approach was implemented, following the methodology proposed by Jing, Chen, Zhang, Li and Zheng (2013) and extended by Thomas, Rane and Harris (2024). The simulation confirmed the robustness of the ranking patterns obtained, showing minimal dispersion and no reversals in task prioritization. These findings reinforce the reliability of the AHP framework even in the presence of subjective variability, enhancing its methodological rigor. While no comparative analysis was performed with alternative methods such as Fuzzy-AHP, TOPSIS, or machine learning techniques, the choice of the AHP was based on its suitability for capturing structured expert knowledge and its capacity to ensure traceability, interpretability, and consistency assessment—features particularly relevant for explaining task-level risk in the absence of large historical datasets. Future studies may benefit from triangulating results using alternative multi-criteria decision-making approaches to expand the methodological landscape and test convergence across models.

6. Conclusions

This methodological article demonstrated the usefulness of the AHP for assessing the automation risk of professional tasks based on structured expert judgment. By decomposing a complex decision problem into hierarchical levels, the proposed approach facilitated conceptually clear comparisons among criteria and alternatives, even in highly complex contexts. The design of didactic and comparative data collection instruments was instrumental in ensuring that experts fully understood the evaluation process, thereby strengthening the validity of their contributions.

The operationalization and validation of the methodology within the accounting profession revealed that, according to aggregated expert judgment using the geometric mean, Record Keeping is the task most susceptible to automation, followed by Reporting. Conversely, tasks involving Data Analysis and Management were considered less vulnerable. This ranking is consistent with prior evidence that routine, standardized activities face a higher risk of technological substitution.

The study also underscores the value of integrating both global and context-specific perspectives, as the diversity of expert backgrounds enriched the analysis and provided a nuanced understanding of discrepancies. Extending this methodology to other professional domains and geographical settings, and incorporating economic, organizational, and regulatory factors, could significantly enhance its explanatory power.

By combining structured expert judgment with multi-criteria decision-making techniques such as AHP, this approach offers a robust and adaptable framework for supporting strategic decisions in the context of the ongoing digital transformation of work

REFERENCES

Acemoglu, D., & Restrepo, P. (2018, January). Artificial intelligence, automation and work. (NBER Working Paper, no. 24196). National Bureau of Economic Research. https://ssrn.com/abstract=3101994

Aguirre, S., & Rodríguez, A. (2017). Automation of a business process using robotic process automation (RPA): A case study. *Proceedings of the Workshop on Engineering Applications*, Cham, Switzerland: Springer, 65–71.

Alaimo, V., Chaves, M.N., & Soler, N. (2019). *El futuro del trabajo en América Latina y el Caribe: ¿Cómo garantizar los derechos de los trabajadores en la era digital?* Banco Interamericano de Desarrollo. https://doi.org/10.18235/0002088

Ali, Y., & Kaleem, M. (2025). Resilience enhancement of tourism supply chain amid uncertain situations: Developing countries, a case in point. *International Journal of the Analytic Hierarchy Process*, *16*(3), 1–18. https://doi.org/10.13033/ijahp.v16i3.1280

Alonso, J., & Lamata, M. (2006). Consistency in the analytic hierarchy process: A new approach. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *14*(4), 445–459. https://doi.org10.1142/S0218488506004114

Arízaga Piedra, T.K., Monge Loja, D.A., & Muñoz Pauta, M.A. (2023). Adaptación Lingüística de la Escala de Desarrollo Armónico 0–5 años. *Revista Cubana de Educación Superior*, *38*(especial).

Arntz, M., Gregory, T., & Zierahn, U. (2016). *The risk of automation for jobs in OECD countries: A comparative analysis*. (OECD Social, Employment and Migration Working Papers, no. 189). OECD Publishing. https://doi.org/10.1787/5jlz9h56dvq7-en

Anderson, D., Sweeny, D., & Williams, T. (2016). *Métodos cuantitativos para los negocios*, 13th ed. Cengage Learning.

Autor, D.H., Levy, F., & Murnane, R.J. (2003). The skill content of recent technological change: An empirical exploration. *Quarterly Journal of Economics*, *118*(4), 1279–1333. https://doi.org/10.1162/003355303322552801

Autor, D.H. (2015). Why are there still so many jobs? The history and future of workplace automation, *Journal of Economic Perspectives*, *29*(3), 3–30. https://doi.org/10.1257/jep.29.3.3

Autor, D.H., & Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market, *American Economic Review*, *103* (5), 1553–1597. https://doi.org/10.1257/aer.103.5.1553

Bannò, E. Filippi, E., &Trento, S. (2021). Rischi di automazione delle occupazioni: Una stima per l'Italia. *Stato e Mercato*, *3*, 315–350. https://doi.org/10.1425/103268

Barragán, S., González, L., & Calderón, G. (2022). Modelling student dropout risk using survival analysis and analytic hierarchy process for an undergraduate accounting program. *Interchange*, *53*, 407–427. https://doi.org/10.1007/s10780-022-09463-7

Barragán, S., & Guzmán, A. (2025). La ruta para revisiones paraguas en ciencias sociales: Explorando nuevos horizontes. *European Public and Social Innovation Review*, *10*, 1–19. https://doi.org/10.31637/epsir-2025-689

Bessen, J.E. (2018). *AI and jobs: The role of demand*. (NBER Working Paper, no. 24235). National Bureau of Economic Research. https://doi.org/10.3386/w24235

Budiño, G., & Asuaga, C. (2022). Características de las tareas de la práctica profesional contable que pueden ser afectadas por la automatización de procesos: Validación de metodología y análisis en un caso de actividad gerencial. *Revista Gestión y Desarrollo Libre*, *7*(14), 1–13. https://doi.org/10.18041/2539-3669/gestionlibre.14.2022.9379

Buzzelli, G. (2023). Automation, globalization and the (mis)perception of risks: New evidence from six advanced economies. *Rivista Italiana di Politiche Pubbliche*, *18*(2), 269–310. https://www.rivisteweb.it/download/article/10.1483/108266

Chambergo, D., Díaz, M. & Benites, V. (2021). Revisiones de alcance, revisiones paraguas y síntesis enfocada en revisión de mapas: Aspectos metodológicos y aplicaciones. *Revista Peruana de Medicina Experimental y Salud Pública*, *38*(1), 136–142. https://doi.org/10.17843/rpmesp.2021.381.6501

Davoodi, A. (2009). On inconsistency of a pairwise comparison matrix. *International Journal of Industrial Mathematics*, *1*, 343–350. https://www.magiran.com/p882888

Dicks, A., Künn-Nelen, A., Levels, M., & Montizaan, R. (2024). Automation risks of vocational training programs and early careers in the Netherlands. *Acta Sociologica*, *67*(4), 507–528. https://doi.org/10.1177/00016993241250277

Filippi, E., Bannò, M., & Trento, S. (2023). Automation technologies and their impact on employment: A review, synthesis and future research agenda. *Technological Forecasting and Social Change*, *191*, 122448. https://doi.org/10.1016/j.techfore.2023.122448

Frey, C.B., & Osborne, M.A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change*, *114*, 254–280. https://doi.org/10.1016/j.techfore.2016.08.019

Goos, M., Manning, A., & Salomons, A. (2009). Job polarization in Europe, *American Economic Review*, *99*(2), 58–63. https://doi.org/10.1257/aer.99.2.58

Grant M., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, *26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Jing, L., Chen, B. Zhang, B., Li, P., & Zheng, J. (2013). Monte Carlo simulation–aided Analytic Hierarchy Process approach: Case study of assessing preferred non-point-source pollution control best management practices. *Journal of Environmental Engineering*, *139*(5), 618–626. https://doi.org/10.1061/(ASCE)EE.1943-7870.0000673

Im, Z.J., & Schneider, S.M. (2022). Dealing with technological change: Social policy preferences and institutional context. Journal of European Public Policy, *29*(5), 710–731. https://doi.org/10.1080/13501763.2021.1876046

Kendall, M. (1938). A new measure of rank correlation. *Biometrika*, *30*(1–2), 81–89. https://doi.org/10.1093/biomet/30.1-2.81

Kraus, S., Breier, M., Lim, W., Dabić, M., Kumar, S., Kanbach, D., Mukherjee, D., Corvello, V., Piñeiro-Chousa, J., Liguori, E., Palacios-Marqués, D., Schiavone, F., Ferraris, A., Fernances, C., & Ferreira, J. (2022). Literature reviews as independent studies: guidelines for academic practice. *Review of Managerial Science*, *16*, 2577–2595. https://doi.org/10.1007/s11846-022-00588-8

Latham, S., & Humberd, B. (2018). Four ways jobs will respond to automation. *MIT Sloan Management Review*, *60*(1), 11–14. https://sloanreview.mit.edu/article/four-ways-jobs-will-respond-to-automation/

Landriscina, F. (2013). *Simulation and learning*. Springer. https://doi.org/10.1007/978-1-4614-1954-9

Lorenz, H., Stephany, F., & Kluge, J. (2023). The future of employment revisited: How model selection affects digitization risks. *Empirica*, *50,* 323–350. https://doi.org/10.1007/s10663-023-09571-2

Mantogiannis, V.A., & Katsigiannis, F.A. (2020). Assessing real estate investment alternatives: A multi-criteria and multi-stakeholder decision aid tool. *International Journal of the Analytic Hierarchy Process*, *12*(1), 136–165. https://doi.org/10.13033/ijahp.v12i1.702

Maskavizan, A.J., Poco, A.N., & Calzolari, A. (2023). Aspectos prácticos del uso del coeficiente de concordancia W de Kendall para el jueceo de cuestionarios en enfermería. *Arandu Poty*, *2*(2), 23–32. https://divulgacioncientifica.unca.edu.py/index.php/AranduPoty/article/download/61/55/284

Mendoza, D. (2013). Aplicación de la teoría de decisión multicriterio discreta para ponderar factores en procesos de acreditación de alta calidad. *INGE CUC*, *9*(1), 25–41.

Munier, N. (2011). *A strategy for using multicriteria analysis in decision-making: A guide for simple and complex environmental projects*. Springer.

Nedelkoska, L. & Quintini, G. (2018, March). *Automation, skills use and training*. (OECD Social, Employment and Migration Working Papers, no. 202). OECD Publishing. https://doi.org/10.1787/2e2f4eea-en

Nguyen, A.T., Tuyen, V. V., Than, V. H., & Tran, V. K. (2025). Applying the analytic hierarchy process to prioritize risks of the food supply chain in Vietnam. *International Journal of the Analytic Hierarchy Process*, *17*(1), 1–33. https://doi.org/10.13033/ijahp.v17i1.1199

Ogosi, A., Lira, J., Guadalupe, V., Lira, Z., & Cacsire, P. (2023). Aplicación de inteligencia artificial para la predicción del estrés financiero en empresas del sector de recogida y tratamiento de aguas residuales. *Revista Ibérica de Sistemas e Tecnologias de Informação*, (E58), 152–160. https://hdl.handle.net/20.500.14308/4799

Organización para la Cooperación y el Desarrollo Económico (OCDE). (2019). *OECD employment outlook 2019: The future of work*. OECD Publishing. https://doi.org/10.1787/9ee00155-en

Organización Internacional del Trabajo (OIT). (2021). Panorama laboral 2021 de América Latina y el Caribe. Geneva, Switzerland: OIT. https://www.ilo.org

Ripani, L., Soler, N., Kugler, A., Kugler, M., & Rodrigo, R. (2020). *El futuro del trabajo en América Latina y el Caribe: ¿Cuál es el impacto de la automatización en el empleo y los salarios?* Banco Interamericano de Desarrollo. https://doi.org/10.18235/0002960

Saaty, T.L. (1980). *The Analytic Hierarchy Process*. McGraw-Hill.

Saaty, T.L. (2008). Decision making with the analytic hierarchy process. *Internationl Journal of Services Sciences*, *1*(1), 83–98. https://doi.org/10.1504/IJSSCI.2008.017590

Schmidpeter, B., & Winter-Ebmer, R. (2021). Automation, unemployment, and the role of labor market training. *European Economic Review*, 137, 103808. https://doi.org/10.1016/j.euroecorev.2021.103808

Thomas, T.A., Rane, S. & Harris, J.T. (2024). A case study assessing the integration of nuclear safety and security in facilities using a Monte Carlo simulation-aided Analytical Hierarchy Process. *International Journal of Nuclear Security*, *9*(1). https://doi.org/10.7290/ijns09256645 van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*, 523–538. https://doi.org/10.1007/s11192-009-0146-3

Włoch, R., Śledziewska, K., & Rożynek, S. (2025). Who's afraid of automation? Examining determinants of fear of automation in six European countries. *Technology in Society, 81,* 102782. https://doi.org/10.1016/j.techsoc.2024.102782