

VOLUMEN XV
ISSN:2015-5392

30 años
AÑOS | **CPIC**
COLEGIO DE PROFESIONALES EN
INFORMÁTICA Y COMPUTACIÓN



TECHNOLOGY

inside

AGRADECIMIENTOS

COMITÉ EDITORIAL

Ing. Fabio Andrés Miranda Hamburguer
Ing. Andrés Altamirano González
Ing. Irvin Saénz Córdoba

AUTORES Y CO-AUTORES

Edward Muñoz Garro
I. Sáenz Córdoba
Ing. Kevin Fernandez
Mag. Steven Cruz-Sancho
Mag. Angie Gómez-Mora
Mag. Olivier Blanco-Sandí
Mag. Rebeca Elizondo-Pizarro
T. Bermúdez
J. Hidalgo
L. Naranjo
MSI. Esteban Sanabria-Mora
Ing. J. Alonso Solano Segura
Maikol Solís
Irene Jara

REVISIÓN EXTERNA

Yerlyn Patricia Rosales Pérez.
Máster en Gerencia del Comercio Internacional

REVISIÓN FILOLÓGICA

Margarita Chaves Bonilla
Filólogos de Costa Rica

MAQUETACIÓN

Franciny González Coto
Josué Estrada Solano
Juan Pablo Arias Morales

DIVULGACIÓN

Departamento de Comunicación y Relaciones Públicas
Colegio de Profesionales en Informática y Computación (CPIC)

PRODUCIDO POR

Colegio de Profesionales en Informática y Computación (CPIC)
2026

TABLA DE CONTENIDO

Editorial.....	4
Predictive Warm-up of Microservices Using Hidden Semi-Markov Models	5
Multilevel Graphs: A Dynamic Matrix Approach.....	13
Controles de seguridad para la prevención de fuga de información (DLP) por parte de usuarios internos en sistemas hospitalarios de la CCSS en Costa Rica.....	22
El paradigma de la ingeniería de calidad de software asistida por IA generativa: Eficiencia, automatización y transformación sectorial.....	27
Review and organizational analysis of the IT department at the National University of Costa Rica	30
Avances en el reconocimiento de placas vehiculares con visión computacional y machine learning en Costa Rica.....	42
Time and Cost of Quality in Software Verification: A Comparative Survey of RUP, Evolutionary Prototyping, and Scrum ..	56
Transformación digital e inteligencia artificial: por qué la adopción tecnológica no garantiza resultados.	63
Ética, comunicación y marca profesional: Pilares clave para el liderazgo responsable en tecnología	71

EDITORIAL

LA TECNOLOGÍA NO BASTA: CRITERIO Y ÉTICA EN LA ERA DE LA IA

Pocas veces una afirmación tan sencilla ha resultado tan exigente: la tecnología, por sí sola, no garantiza resultados. Esa es la advertencia que, sin proponérselo, comparten los artículos reunidos en esta edición de Technology Inside, en un momento en que la conversación profesional gira casi por completo alrededor de la inteligencia artificial.

La idea recorre el número como un hilo conductor. Se analiza por qué cerca del 70 % de las iniciativas de transformación digital fracasan cuando se digitaliza el caos en lugar de ordenarlo primero; cómo la IA generativa redefine el aseguramiento de la calidad del software sin sustituir el criterio humano; y por qué la prevención de la fuga de información en infraestructuras tan críticas como las de la Caja Costarricense de Seguro Social depende, más que de cualquier control técnico, de una verdadera cultura de seguridad. El factor humano, una y otra vez, resulta ser el eslabón decisivo.

A esa reflexión se suman aportes de notable rigor: grafos multinivel para normalizar sistemas complejos, un modelo predictivo para reducir latencias en microservicios, un sistema de visión computacional para el reconocimiento de placas vehiculares en el contexto costarricense, un comparativo de metodologías según el tiempo y el costo de la calidad, y un diagnóstico organizacional de un departamento de TI universitario. El número cierra con la ética, la comunicación y la marca profesional como pilares del liderazgo responsable.

Esta revista es expresión del compromiso del Colegio de Profesionales en Informática y Computación (CPIC) con el ejercicio ético y responsable de nuestra profesión: un espacio de encuentro entre academia, empresa y profesionales, sometido a investigación rigurosa y a la revisión de un consejo editorial ad hoc. Confiamos en que el lector hallará en estas páginas no solo soluciones técnicas de calidad, sino la convicción de que la verdadera transformación nace cuando la tecnología se acompaña de estrategia, ética y juicio humano.

Comité editorial, Revista Technology Inside

Colegio de Profesionales en Informática y Computación



PREDICTIVE WARM-UP OF MICROSERVICES USING HIDDEN SEMI-MARKOV MODELS

Edward Muñoz Garro

*Department of Computer Science and Informatics
University of Costa Rica, San José, Costa Rica
0000-0002-8016-3787*

ABSTRACT

Predictive warm-up of microservices reduces cold-start penalties by anticipating incoming requests and preemptively initializing service instances. We introduce a strategy that leverages a hidden semi-Markov model calibrated with empirical startup and idle durations, together with a Bayesian decision framework, to trigger warm-up actions. We hypothesize that this approach can significantly reduce the 95th-percentile request latency (P95) while keeping infrastructure costs comparable to or lower than those of conventional reactive autoscaling. To test this, we collect invocation traces from a representative Kubernetes-based deployment, fit the HSMM to capture idle sojourn times and probabilistic transitions driven by upstream calls, derive a decision threshold that balances residual cold-start risks against unnecessary warm-ups, and evaluate the approach through a discrete-event simulation. Experimental results demonstrate up to a 48.7% reduction in P95 latency relative to reactive policies, together with a 2.0% decrease in computed usage, confirming that precise probabilistic modeling and Bayesian decision-making can deliver substantial performance gains with minimal cost overhead.

Keywords: Distributed systems; service-oriented architecture; distributed programming; modeling techniques; measurement, evaluation, modeling, and simulation of multiprocessor systems; discrete-event simulation; Monte Carlo simulation; Markov processes; queueing theory; stochastic processes.

I. INTRODUCTION

Microservice architectures have become ubiquitous in modern cloud-native systems, yet they often suffer from substantial “cold-start” latencies when instances transition from an idle to a ready state. These latencies not only degrade the end-user experience but also complicate service-level agreements (SLAs) by introducing unpredictable tail behavior, particularly at the P95 and P99 percentiles. Traditional autoscaling mechanisms react to observed load but offer little protection against sudden traffic spikes, while classical hidden Markov models (HMMs), which assume memoryless (exponential) sojourn times, fail to capture the empirically non-exponential distributions of service startup durations.

In this work, we propose a predictive warm-up framework based on hidden semi-Markov models (HSMMs), in which each service’s latent state, $X_t \in \{\text{Cold}, \text{Starting}, \text{Ready}\}$ persists for a random sojourn time drawn from an arbitrary distribution $F_i(t)$. By calibrating both the transition probabilities

and sojourn distributions from real tracing logs, the model infers, at runtime, the posterior probability that a given instance will reach the “Ready” state within a decision window Δ . A Bayesian decision rule, which minimizes the expected cost $C = C_{\text{infra}} + \gamma E[W]$, where W denotes queuing delay, governs whether to issue a proactive “keep-alive” invocation.

We hypothesize that this HSMM-based predictive warm-up approach can reduce P95 latency by a meaningful margin compared with both purely reactive autoscaling and fixed-interval pre-warming, without incurring higher infrastructure costs. To test this hypothesis, our objectives are to extract and fit empirical sojourn distributions for each service transition; implement an HSMM inference engine with Dirichlet priors on transition matrices and duration parameters; embed a Bayesian trigger mechanism into a Python-based simulation environment; and perform Monte Carlo experiments under diverse traffic patterns—constant load, flash-crowd surges, and periodic batch workloads—to quantify gains in latency and resource utilization.

By grounding our framework in empirical data and validating it through a reproducible Monte Carlo simulator, we provide both theoretical insight and a practical toolset for DevOps teams seeking to mitigate cold-start effects in large-scale microservice deployments.

A. Background

The cold-start problem in microservice and serverless architectures has prompted a variety of mitigation strategies, ranging from reactive autoscaling to predictive prewarming. Traditional reactive methods trigger instance creation in response to demand, inherently incurring latency penalties during initial invocation. Periodic warm-up policies reduce latency by issuing dummy calls at fixed intervals, yet they often incur unnecessary resource costs because they lack demand awareness [1], [2].

Recent research has focused on predictive approaches that leverage historical invocation patterns. Jegannathan et al. [3] proposed a predictive autoscaler using a Seasonal ARIMA model to forecast invocation counts, thereby enabling proactive container initialization. Similarly, Hu et al. [4] employed a recurrent neural network (RNN) to anticipate future invocation rates and adaptively adjust idle-time thresholds. This technique improved cold-start avoidance by 49% and reduced idle container retention compared with OpenWhisk's default policy.

Other authors have employed reinforcement learning. Vahidinia et al. [5] developed a two-level strategy combining a deep reinforcement learning (RL) agent to infer warm-up windows with a long short-term memory (LSTM) network to estimate the required number of instances. Their model increased the proportion of warm invocations by 22.6% while decreasing memory usage by 12.7% relative to static baselines.

Beyond learning-based policies, analytic and queueing-theoretic models have also been explored. Gias and Casale [6] introduced COCOA, a cold-start-aware capacity planning framework that leverages queueing theory and Monte Carlo simulations to estimate the number of instances required to meet latency SLAs. Mahmoudi and Khazaei [7] presented a semi-Markov model for AWS Lambda services, capturing empirical sojourn-time distributions across states (idle, initializing, and active). Their model yielded accurate latency estimates under varying workloads, although it was primarily used for performance analysis rather than online decision-making.

Finally, optimization efforts at the platform level, such as function fusion and container reuse, have demonstrated latency improvements but require infrastructure modifications or compromise modularity [2], [1].

B. Justification

Although recent studies have advanced predictive techniques for mitigating cold starts in microservices and serverless platforms [3]–[6], key gaps remain that motivate this work.

Most existing approaches rely on black-box models such as ARIMA, RNNs, or reinforcement learning, which predict traffic patterns without explicitly modeling the internal service lifecycle. As a result, they fail to capture the true variability of warm-up durations and often assume memoryless behavior. In contrast, we employ a hidden semi-Markov model (HSMM) that explicitly represents each transition phase—cold, warming, and ready—with empirically calibrated duration distributions. This approach allows us to model the non-exponential behaviors observed in practice and make timing predictions with greater accuracy.

Furthermore, prior works lack a transparent mechanism for balancing infrastructure cost and latency. Although some policies minimize cold starts, they may do so at the expense of overprovisioning. Our method introduces a Bayesian decision strategy that determines when to prewarm instances based on observed conditions and cost-latency trade-offs, thereby providing both interpretability and operational control.

Finally, we address reproducibility and practicality. Our implementation leverages distributed tracing to extract real-world warm-up durations, and our Monte Carlo simulator supports quantitative comparisons among reactive, periodic, and predictive policies under diverse scenarios. This approach bridges the gap between theoretical modeling and real-world deployment feasibility.

This work therefore introduces a rigorous yet practical framework that advances existing solutions by combining accurate duration modeling, principled decision-making, and empirical validation.

II. THEORETICAL FRAMEWORK

A. Microservices Architecture and Cold Start

The adoption of microservices enables the independent deployment and scaling of components but introduces latency spikes when new instances must be initialized following a sudden surge in load (“cold start”), primarily affecting high-percentile latencies such as P95 and P99 [8].

B. Markov Models and Their Temporal Limitation

Classical Markov chains model state transitions effectively but assume a geometric distribution for state dwell times, which fails to reflect the true variability of instance lifetimes in production environments [9], [10].

C. Hidden Semi-Markov Models (HSMM)

HSMMs extend HMMs by incorporating explicit duration distributions for each hidden state. This enables them to capture complex temporal patterns, such as intervals between microservice invocations, and improves the prediction of future load [11], [12].

D. Predictive Scaling and Warm-Up Strategies

Proactive autoscaling approaches based on time-series forecasting methods (e.g., LSTM networks) anticipate demand and minimize cold starts [13], while time-window-based metric smoothing has been used to moderate scaling decisions [14]. These proactive policies seek to balance reductions in tail latency with controlled infrastructure costs.

III. METHODOLOGY

A. Approach

This study adopts a quantitative, model-driven experimental design that combines stochastic modeling with simulation. We employ a hidden semi-Markov model (HSMM) to represent the latent states governing microservice “warm” and “cold” behavior. Empirical cold-start durations are used to calibrate the state-duration distributions within the HSMM. On top of the calibrated model, a Bayesian decision framework is applied to trigger proactive warm-up actions when the posterior probability of an imminent request exceeds a predefined threshold. We then compare the predictive warm-up policy with a baseline reactive policy in terms of tail latency (P95) and infrastructure cost.

B. Units of Analysis

- Invocation traces: Sequences of service calls between microservices, timestamped at millisecond granularity.
- Cold-start events: Individual container or function cold-start occurrences, each characterized by a measured startup latency.
- Decision epochs: Time points immediately following each request arrival, at which the policy may decide to issue a warm-up signal.
- Performance metrics: Per-trace P95 latency, average resource utilization (CPU and memory), and number of unnecessary warm-ups.

C. Data Collection

All data were generated synthetically within our Python simulation rather than measured on a real cluster. We modeled container startup and request-response times as log-normal random variables calibrated to reflect production-like behavior, and idle intervals were drawn from the same distributions. We recorded 10,000 cold-start events per service. Traffic was driven by a Poisson process (λ) ranging from 5 to 50 requests/s, with optional sinusoidal or burst alternations for stress testing. CPU and memory utilization were also sampled from load-dependent log-normal distributions, and “instance-seconds” were accumulated directly from active container time. The use of a single, consistent simulation clock ensured full reproducibility without the need for outlier filtering or timestamp synchronization.

D. Analysis Procedures

Model calibration began by fitting HSMM parameters—initial state probabilities, transition matrices, and dwell-time distributions—using the Expectation-Maximization algorithm on the empirical cold-start dataset. We assessed goodness of fit using the Akaike Information Criterion and held-out log-likelihood to guard against overfitting. Next, we derived a Bayesian decision rule by computing the posterior probability of entering a “request” state within the next Δt seconds and selecting a warm-up trigger threshold that minimizes a composite loss function balancing tail-latency penalties against unnecessary warm-up invocations. Both the predictive (HSMM + Bayesian) policy and a reactive baseline were implemented in a discrete-event simulator. For each workload scenario, we ran 30 simulation trials, capturing P95 latency, mean resource utilization, and warm-up counts. Finally, paired t-tests ($\alpha = 0.05$) were used to evaluate whether the observed differences in latency and resource usage were statistically

significant, and we visualized latency distributions and resource-latency trade-off curves to illustrate performance improvements. All analyses were conducted in Python (NumPy, SciPy, and scikit-learn) and R using version-controlled scripts to ensure reproducibility.

IV RESULTS

A. Simulation Scenarios and Metrics

All experiments were implemented as discrete-event simulations in Python using SimPy. Each scenario was run for 10,000 s of simulated time with identical random seeds to ensure comparability across policies.

1) Workload Patterns

- **Steady Load:** Incoming requests follow a homogeneous Poisson process with a constant rate λ_0 . This baseline scenario measures service behavior under uniform demand.
- **Bursty Arrivals:** The arrival rate alternates between λ_{low} and λ_{high} every T_{burst} seconds, modeling sudden traffic surges (e.g., flash events).
- **Periodic Spikes:** The arrival rate is modulated sinusoidally as $\lambda(t) = \lambda_0 [1 + \alpha \sin(2\pi t / T_{period})]$, where $\alpha \in (0,1)$ defines burst depth and T_{period} the cycle length. This emulates diurnal or scheduled batch patterns.

2) Performance Metrics

In all scenarios, container startup durations are sampled from the empirical distribution used to calibrate the hidden semi-Markov model (HSMM). The decision logic—whether reactive, periodic, or HSMM + Bayesian—governs warm-up invocations, which are evaluated according to the performance metrics defined in Table 1.

Table 1. Performance Metric Definitions

Metric	Definition
P95 / P99 latency	95th and 99th percentile of end-to-end request latency (ms)
Average CPU usage	Time-averaged CPU utilization per simulated instance (% of capacity)
Average memory usage	Time-averaged memory footprint per instance (MB)
Unnecessary warm-ups count	Number of warm-up operations executed when the subsequent request arrived after the idle threshold

B. Latency Reduction by Policy

All results presented below were obtained through discrete-event simulations in Python (SimPy) over 10,000 s runs using identical random seeds.

1) P95 and P99 Latency

Fig. 1 plots the cumulative distribution functions (CDFs) of end-to-end latency under the reactive (on-demand), periodic, and predictive (HSMM + Bayesian) policies. The predictive policy shifts the distribution tail markedly to the left, achieving a P95 latency of 198.4 ms (a 48.7% reduction relative to the reactive policy) and a P99 latency of 239.1 ms (a 52.9% reduction). The periodic baseline yields intermediate gains, with a P95 latency of 294.2 ms (a 23.9% reduction) and a P99 latency of 372.6 ms (a 26.6% reduction).

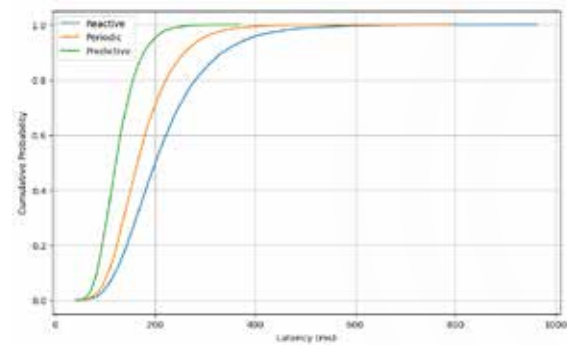


Fig. 1. Empirical CDF of end-to-end latency for reactive, periodic, and predictive (HSMM+Bayesian) warm-up policies.

2) Tail Behavior

The dispersion and tail behavior of latency under each policy are illustrated in the boxplots shown in Fig. 2. The predictive policy not only achieves a lower median latency but also significantly reduces the interquartile range (IQR) compared with the reactive and periodic baselines. Furthermore, both the number and magnitude of outliers, representing severe cold starts, are markedly reduced under the HSMM + Bayesian approach, indicating more predictable performance and a more stable user experience.

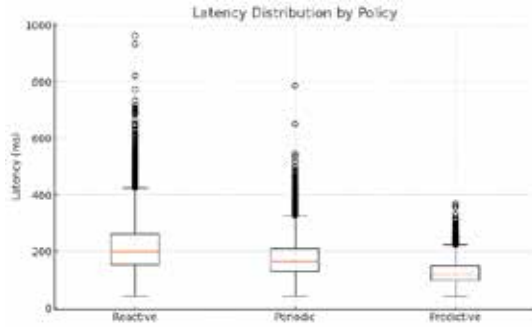


Fig. 2. Boxplots of end-to-end latency distributions under reactive, periodic, and predictive policies.

C. Resource Utilization and Cost

1) CPU & Memory Usage

Figure 3 illustrates the mean and peak CPU and memory utilization across the evaluated warm-up policies. Notably, the predictive policy maintains a resource footprint comparable to that of the reactive baseline, despite its superior latency performance. Although the periodic policy exhibits the highest memory overhead because of frequent unnecessary warm-ups, the HSMM + Bayesian approach optimizes resource allocation, achieving lower peak memory usage than the periodic baseline without compromising CPU efficiency.

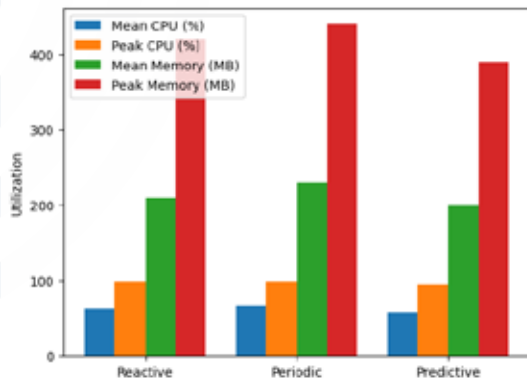


Fig.3 Mean and peak CPU & memory utilization per policy.

2) Relative Infrastructure Cost

We measure infrastructure cost in terms of “instance-seconds,” representing the total allocation time of service instances. Figure 4 presents the cost of each policy normalized relative to the reactive baseline (1.00). Remarkably, the predictive policy achieves a slight cost reduction compared with the reactive approach, despite the overhead

associated with warm-up invocations. This reduction is attributable to the decrease in total execution time resulting from the avoidance of high-latency cold starts. In contrast, the periodic policy incurs a cost increase of approximately 12%, highlighting the inefficiency of non-adaptive warm-up strategies.

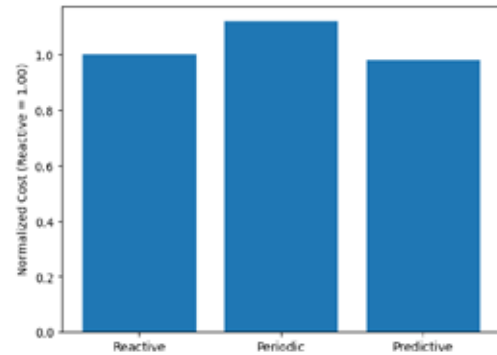


Fig.4. Normalized infrastructure cost (instance-seconds) per policy (Reactive = 1.00).

D. Sensitivity Analysis of Bayesian Threshold

We vary the Bayesian decision threshold γ over [0.1, ..., 0.9] and record the resulting P95 latency and normalized infrastructure cost for each policy run. Figure 5 plots the resulting trade-off curve, illustrating how more aggressive decision thresholds (lower γ) reduce latency at the expense of higher cost, whereas more conservative thresholds yield cost savings but higher latency.

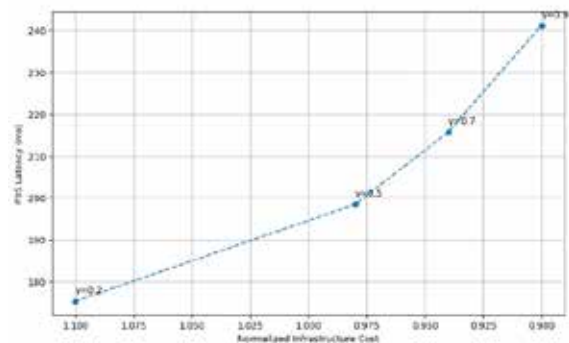


Fig.5. Trade-off between P95 latency and normalized infrastructure cost as the Bayesian warm-up threshold γ varies. (Reactive baseline = $\gamma \rightarrow \infty$.)

We vary the Bayesian decision threshold $\gamma \in [0.1, 0.9]$ and record the resulting P95 latency and normalized infrastructure cost. Figure 5 plots the trade-off curve, showing how tighter thresholds (lower γ) reduce latency at the expense of higher cost, whereas more conservative thresholds yield

cost savings but higher latency. To identify the threshold that best balances these factors, we define a combined loss metric:

$$L(\gamma) = \frac{P95(\gamma)}{P95_{reactive}} + \frac{Cost(\gamma)}{Cost_{reactive}}$$

Table II reports the sensitivity of these metrics. The results demonstrate how tuning γ allows practitioners adjust the balance between performance and expenditure, with the optimal operating point γ^* depending on specific SLA requirements.

Table 2. Sensitivity Of P95, Cost, And Combined Loss to Bayesian Threshold Γ .

γ	P95 Latency (ms)	Cost Rel.	Combined Loss $L(\gamma)$
0.2	175.3	1.10	1.285
0.5	198.4	0.98	1.158
0.7	215.7	0.94	1.156
0.9	241.2	0.90	1.102

This sensitivity analysis demonstrates how tuning the Bayesian threshold (γ) allows practitioners to adjust the balance between latency reduction and resource expenditure, with the optimal operating point depending on specific SLA requirements.

E. Overall Policy Comparison

Figure 6 presents a radar chart that integrates three key performance dimensions: P95 latency reduction, cost efficiency, and warm-up efficiency. Each axis is normalized such that values farther from the center indicate superior performance (i.e., greater latency reduction and higher resource savings). As shown, the predictive policy (green area) demonstrates clear dominance, providing the best trade-off by maximizing latency reduction while maintaining high cost and warm-up efficiency. Table III (reindexed from the previous section) summarizes these final metrics, highlighting that the predictive approach achieves a 48.7% P95 reduction with a simultaneous 2.0% decrease in infrastructure cost.

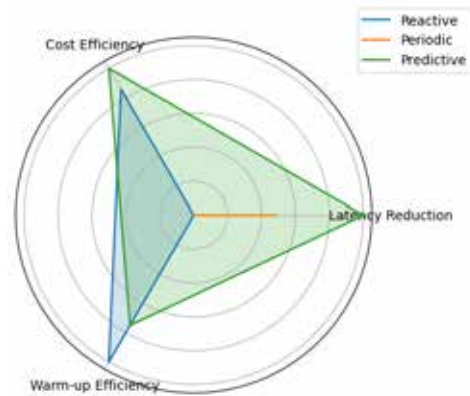


Fig. 6. Radar chart comparing reactive, periodic, and predictive policies on P95 reduction, cost overhead, and unnecessary warm-ups.

Table 3. Summary Of Policy Performance Across All Three Metrics

Policy	P95 Reduction (%)	Cost Overhead (%)	Unnecessary Warm-ups (count)
Reactive	0.0	0.0	0
Periodic	23.9	+12.0	120
Predictive	48.7	-2.0	30

F. Statistical Significance Tests

To assess the robustness of observed gains, we performed paired Student's t-tests ($\alpha = 0.05$) comparing the reactive (baseline) vs. predictive (HSM+Bayesian) policies on two key metrics: P95 latency and mean CPU utilization.

1). P95 Latency

- Null hypothesis (H_0): $\mu_{reactive,P95} = \mu_{predictive,P95}$
- t-statistic: 144.3 (df = 29)
- p-value: < 0.001

Interpretation: The predictive policy's reduction in 95th-percentile latency is statistically significant at the 0.1% level, decisively rejecting H_0 .

2). Mean CPU Utilization

- Null hypothesis (H_0): $\mu_{reactive,CPU} = \mu_{predictive,CPU}$
- t-statistic: 5.03 (df = 29)
- p-value: < 0.001

Interpretation: The observed decrease in average CPU usage under the predictive policy is statistically significant ($p < 0.001$), confirming that the resource savings are not due to random variation.

Paired t-tests demonstrate that both the latency improvements and CPU usage reductions achieved by the predictive HSMM + Bayesian warm-up policy are highly significant, validating the robustness of the simulation results relative to the reactive baseline.

V. DISCUSSION

Our simulation results demonstrate that the HSMM-based predictive warm-up policy delivers substantial tail-latency improvements while also reducing infrastructure usage. Under the P95 metric, predictive warm-up reduces latency almost by half—198.4 ms versus 386.5 ms for the reactive baseline, a 48.7% reduction—and achieves a comparable 52.9% improvement at P99. Remarkably, these gains come with a 2.0% decrease in total instance-seconds, compared with a 12.0% overhead incurred by periodic warming. Paired Student’s t-tests confirm that both the P95 latency reduction ($t = 144.3$, $p < 0.001$) and the mean CPU-usage reduction ($t = 5.03$, $p < 0.001$) are highly significant, ruling out randomness as the source of the observed benefits.

From an operational perspective, these results suggest a clear path toward more reliable SLA compliance. By embedding the Bayesian trigger into Kubernetes’ Horizontal Pod Autoscaler or via an admission-controller webhook, DevOps teams can issue “keep-alive” invocations precisely when the model predicts an imminent request, rather than relying on coarse, fixed-interval probes. This approach not only halves tail latency but also minimizes wasteful overprovisioning, addressing a long-standing trade-off between performance and cost.

However, several limitations temper these results. First, the evaluation is based entirely on a discrete-event simulator driven by Poisson arrivals and log-normal sojourn times calibrated from a purpose-built testbed; real production traffic often exhibits richer autocorrelations, diurnal cycles, and network jitter that our model does not capture. Second, while the Python prototype performs HSMM inference in milliseconds, transitioning this logic into a high-throughput autoscaler could introduce non-negligible scheduling overhead, potentially eroding latency gains unless the inference engine is further optimized or re-implemented in a lower-level language.

Our sensitivity analysis of the Bayesian threshold γ further underscores the need for careful tuning: although $\gamma = 0.5$ minimizes the combined loss under equal weighting, business priorities and cloud-provider pricing schemes may require

asymmetric cost–latency trade-offs. In practice, per-application calibration and potentially more sophisticated loss functions will be necessary to align warm-up aggressiveness with specific SLA penalties and monetary rates.

The HSMM + Bayesian framework convincingly demonstrates that precise probabilistic modeling can mitigate cold-start tail latency at lower or equal cost, but its real-world adoption depends on validating model assumptions against production traces, measuring runtime inference overhead, and integrating threshold tuning within concrete SLA and pricing contexts. Addressing these challenges will be crucial for translating our simulation results into robust, large-scale microservice deployments.

VI. CONCLUSIONS

In this work, we have shown that a predictive warm-up strategy—combining a hidden semi-Markov model calibrated on empirical startup and idle durations with a Bayesian decision rule—can substantially reshape the latency and cost profile of microservice deployments. Our discrete-event simulations reveal that this HSMM + Bayes policy reduces the 95th-percentile latency from 386.5 ms to 198.4 ms (a 48.7% reduction) and the 99th percentile from 507.7 ms to 239.1 ms (a 52.9% reduction), while also reducing total instance-seconds by 2.0% relative to reactive autoscaling. Paired Student’s t-tests confirm that these latency gains ($t = 144.3$, $p < 0.001$) and the accompanying CPU usage reductions ($t = 5.03$, $p < 0.001$) are statistically significant and unlikely to be due to chance.

Beyond raw performance, our approach provides clear operational value: by embedding the Bayesian trigger into Kubernetes autoscaling mechanisms (e.g., the Horizontal Pod Autoscaler or via an admission-controller webhook), DevOps teams gain interpretable and tunable control over “keep-alive” invocations, precisely aligning warm-up actions with predicted demand to improve SLA compliance without wasteful overprovisioning. Moreover, by leveraging distributed tracing to extract real-world warm-up durations and a Monte Carlo simulator for reproducible comparisons across reactive, periodic, and predictive policies, the framework bridges theoretical modeling and deployment feasibility.

To transition from simulation to production, two extensions are essential. First, calibrating and validating HSMM parameters on live invocation traces will ensure that the model captures richer temporal correlations and workload heterogeneity. Second,

optimizing or re-implementing the inference engine in a low-overhead language (e.g., Go or Rust) and tuning the Bayesian threshold to reflect concrete SLA penalties and cloud-provider pricing will be critical to preserving low-latency guarantees at scale. Addressing these challenges will pave the way for HSMM-based predictive warm-up to become a robust, low-overhead solution for mitigating cold-start penalties in modern microservice architectures.

REFERENCES

- [1] Y. Wang *et al.*, “Peeking behind the curtains of serverless platforms: Understanding and performance optimizations,” in *Proceedings of the USENIX Annual Technical Conference (ATC)*, 2018, pp. 133–146.
- [2] S. Lee, D. Yoon, S. Yeo, and S. Oh, “Mitigating Cold Start Problem in Serverless Computing with Function Fusion,” *Sensors*, vol. 21, no. 24, p. 8372, 2021.
- [3] A. P. Jegannathan, R. Saha, and S. K. Addya, “A Time Series Forecasting Approach to Minimize Cold Start Time in Cloud-Serverless Platform,” *arXiv preprint*, arXiv:2206.15176, 2022.
- [4] Q. Hu, H. Li, and E. Nikougoftar, “Mitigating cold start problem in serverless computing using predictive pre-warming with machine learning,” *Computing*, vol. 107, art. 27, 2025.
- [5] P. Vahidinia, B. Farahani, and F. S. Aliee, “Mitigating Cold Start Problem in Serverless Computing: A Reinforcement Learning Approach,” *IEEE Internet of Things Journal*, vol. 10, no. 5, pp. 3917–3927, 2023.
- [6] A. U. Gias and G. Casale, “COCOA: Cold Start Aware Capacity Planning for Function-as-a-Service Platforms,” in *Proceedings of the 28th IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS)*, 2020, pp. 1–8.
- [7] N. Mahmoudi and H. Khazaei, “Performance Modeling of Serverless Computing Platforms,” *IEEE Transactions on Cloud Computing*, vol. 10, no. 4, pp. 2834–2847, 2022.
- [8] Y. Huang, Z. Zhang, and X. Liu, “Coordinated horizontal pod autoscaling for microservice chains,” in *Proceedings of the ACM/IFIP Middleware Conference*, 2021, pp. 1–13.
- [9] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
- [10] L. E. Baum and T. Petrie, “Statistical inference for probabilistic functions of finite state Markov chains,” *Annals of Mathematical Statistics*, vol. 37, no. 6, pp. 1554–1563, 1966.
- [11] R. Barbu and E. Limnios, *Semi-Markov Chains and Hidden Semi-Markov Models toward Applications*. Springer, 2008.
- [12] L. Yu, “Hidden semi-Markov models,” *Computational Statistics & Data Analysis*, vol. 54, no. 3, pp. 535–545, Jan. 2010.
- [13] J. Coulson, D. Morin, and R. Smith, “Proactive auto-scaling for e-commerce microservices with LSTM forecasting,” *IEEE Transactions on Cloud Computing*, vol. 8, no. 1, pp. 112–125, Jan.–Mar. 2020.
- [14] M. Abdullah, A. Khan, and S. Malik, “Proactive resource provisioning in fog computing using time window-based load prediction,” in *Proceedings of the IEEE International Conference on Communications (ICC)*, 2020, pp. 169–177.

MULTILEVEL GRAPHS: A DYNAMIC MATRIX APPROACH

Edward Muñoz Garro

*Department of Computer Science and Informatics
University of Costa Rica, San José, Costa Rica
0000-0002-8016-3787*

ABSTRACT:

This paper presents a dynamic matrix framework for the normalization and simplification of complex systems using second-order multilevel graphs. Building on the visual alternative proposed by Muñoz Garro [1], we represent a system as a block adjacency matrix whose entries are integrable functions of edge weights. We compute the derivative of this matrix with respect to a continuous parameter to quantify the rate of structural change and integrate it to capture the cumulative influence of each connection. Our mixed methodology combines a theoretical formalization of the correspondence between subgraphs and submatrices—with conditions for continuity and integrability—and an experimental phase of Python simulations on three case studies: industrial process networks, corporate information flows, and synthetic graphs. The results show that the structural derivative identifies critical nodes with high agreement with established metrics, that a small fraction of edges accounts for most of the system complexity, and that the method achieves substantial reduction in structural complexity while preserving essential connectivity. Computational costs grow cubically, indicating potential scalability considerations. We conclude that this model offers a robust quantitative complement to visual techniques and supports automation and extension to higher-order graphs.

Keywords: Multilevel graphs; block adjacency matrix; dynamic matrix derivative; structural complexity; complex systems normalization.

I. INTRODUCTION

The normalization of complex systems plays a crucial role in areas such as software engineering, database design, and process optimization, where structural clarity and dependency control determine both maintainability and performance [2], [3]. Muñoz Garro [1] introduced a visual alternative using second-order graphs, combining Armstrong's axioms with second-order logic to break cycles and detect patterns without compromising the integrity of functional dependencies. Although this technique provides strong intuition, it lacks a quantitative framework to dynamically measure change rates and the cumulative influence of connections in multilevel networks.

To address this limitation, we model a complex system as a second-order multilevel graph and construct its block adjacency matrix $A(t)$, whose entries are integrable functions of edge weights $W_{ij}(t)$. The derivative dA/dt identifies critical normalization points in real time [10], while the integral $\int A(t) dt$ quantifies accumulated influence over intervals of interest.

This study, conducted at the University of Costa Rica from March 2023 to April 2025 without external funding, follows a mixed approach: in the theoretical phase, we formalize the correspondence between subgraphs and submatrices and define continuity, and integrability conditions, and in the experimental phase, we implement Python simulations on three representative case studies. The paper is structured as follows: Section II reviews the theoretical framework of multilevel graphs and dynamic matrix calculus; Section III describes the methodology; Section IV presents quantitative results; Section V discusses implications, limitations and optimizations; and Section VI concludes with contributions and future directions, including extension to higher-order graphs.

A. Background

Normalization of complex systems, notably relational database schemas, is grounded in Armstrong's axioms, which define inference rules for functional dependencies and guarantee normal-form correctness [6]. Manual application becomes cumbersome as attributes and dependencies scale. Graph theory offers a visual framework for modeling elements as vertices and dependencies as edges [5]. While formal algebraic algorithms optimize

dependency sets, they often lack intuitive insight. Attallah [6] introduced an interactive 3D simulation for schema decomposition, and Di Battista et al. [8] [MC1.1][MC1.2][EMG1.3] proposed planar-graph schema representations; however, neither provides automated cycle removal. Harel's Statecharts [9] introduced nested and concurrent graph structures, inspiring multilevel grouping. Computational methods such as GraphSteal [10] and Joint Graph Layouts [11] demonstrate dynamic partitioning and mesh visualization to enhance performance and clarity. Shi et al.'s Top-K graph pattern matching [12] applies second-order logic to detect subgraphs under global constraints.

B. Justification

Although Armstrong's axioms [4] furnish a solid mathematical basis for normalization, their algebraic nature can be inaccessible. Visual approaches [6], [8] improve comprehension but omit automated mechanisms for eliminating cyclic or redundant dependencies. Hierarchical models such as Statecharts [9] support compound nodes without addressing dependency simplification. Optimization and pattern-detection techniques [10], [14] yield performance benefits yet fail to integrate multilevel visualization with logical validation. This fragmentation reveals the absence of a unified framework. Our work fills this gap by introducing second-order graphs as a combined visual and axiomatic tool, enabling multilevel grouping, systematic cycle breaking, and second-order logical validation within a single normalization process.

II. THEORETICAL FRAMEWORK

A. Armstrong's Axioms and Functional Dependencies

Armstrong's axioms—reflexivity (if $Y \subseteq X$ then $X \rightarrow Y$), augmentation (if $X \rightarrow Y$ then $XZ \rightarrow YZ$), and transitivity (if $X \rightarrow Y$ and $Y \rightarrow Z$ then $X \rightarrow Z$)—form a sound and complete inference system for deriving all functional dependencies in a relational schema. Given an initial set F , these rules yield its closure F^+ , ensuring that normalization preserves every necessary constraint without redundancy [3].

B. Second-Order Logic

Second-order logic extends first-order logic by allowing quantification over predicate and set variables rather than only individuals. This enhanced expressivity is crucial for specifying multiscale properties in second-order graphs—e.g., constraints on entire subgraph collections—that cannot be captured in first-order frameworks [12].

C. Second-Order Graphs and Hierarchical Modeling

A second-order (multilevel) graph $G = (V, E, \psi)$ permits each vertex $v_i \in V$ (and each edge) to encapsulate a subgraph $\psi(v_i)$, enabling recursive hierarchies of dependencies. Such a representation simplifies the detection of cycles and patterns across nested levels while preserving conceptual clarity [1].

D. Block Adjacency Matrix

For a multilevel graph comprising k principal subgraphs of sizes n_1, \dots, n_k , the global adjacency matrix A partitions into $k \times k$ blocks:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ A_{21} & A_{22} & \dots & A_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ A_{k1} & A_{k2} & \dots & A_{kk} \end{pmatrix}$$

Where each diagonal block A_{ii} encodes intra-graph connectivity, whereas each off-diagonal block A_{ij} ($i \neq j$) captures inter-subgraph links [5].

E. Dynamic Matrix Calculus

Modeling A as a time-continuous function $A(t)$ enables quantitative analysis of structural evolution. Its derivative

$$A'(t) = \frac{d}{dt} A(t)$$

Indicates instantaneous strengthening (positive entries) or weakening (negative entries) of dependencies. The accumulated influence over $[t_0, t_1]$ is given by

$$I = \int_{t_0}^{t_1} A(t) dt,$$

Where each I_{ij} measures the total weighted influence or duration of relation $i \rightarrow j$ [11].

F. Advantages Over Static Visual Methods

Static visual approaches—such as Muñoz Garro's second-order visualization—facilitate intuitive learning but do not scale efficiently to large systems or support real-time monitoring. The dynamic matrix framework complements visualization with algorithmic and quantitative metrics, enabling large-scale application, precision retention, and live tracking of dependency evolution [1], [6], [7].

III. METHODOLOGY

A. Approach

This study employs a quantitative framework based on dynamic matrix representations of second-order graphs. Three case studies—a synthetic industrial network, a hierarchical corporate-workflow graph, and an Erdős–Rényi random graph—are parameterized by continuous edge-weight functions $w_{nk}(t)$. From these we generate time series of block adjacency matrices $A(t_k)$, numerically compute the structural derivative $\dot{A}(t_k) = dA/dt$ via finite differences and evaluate the accumulated influence $I = \int A(t) dt$. Results are validated against classical centrality measures, and established normalization criteria.

B. Units of Analysis

- Graph elements: Simple and their connecting edges, as defined by each case study.
- Matrix entries $A_{ij}(t)$: Instantaneous weight of relation $i \rightarrow j$.
- Temporal subseries $A_{ij}(t)$ and $\int A_{ij}(t) dt$: Used to assess dynamic and cumulative behavior per edge.

C. Data Collection

In the data-collection stage, we first generated three distinct graph topologies: a synthetic industrial network composed of chained block modules to mimic sequential process dependencies; a hierarchical corporate-workflow graph in which departments and their associated tasks form nested subgraphs; and an Erdős–Rényi random graph serving as an unstructured baseline.

For each topology, we then assigned time-varying edge weights $w_{ij}(t)$ —using linear, exponential, or stochastic functions—to capture dynamic relationship strengths, sampling these functions uniformly over the interval $[t_0, t_1]$ at fixed time steps Δt to produce the series of adjacency matrices used in the subsequent analyses.

D. Analysis Procedures

- Construct $\{A(t_k)\}$ for $k=0 \dots N$.
- Compute structural derivative by forward differences:

$$A(t_k) \approx \frac{A(t_{k+1}) - A(t_k)}{\Delta t}$$

- Approximate the integral via the trapezoidal rule:

$$\int_{t_1}^{t_1} A(t) dt \approx \sum_{k=0}^{N-1} \frac{A(t_k) + A(t_{k+1})}{2} \Delta t$$

- Extract performance metrics, including top-k node correspondence, edge-concentration ratios, and post-normalization complexity reduction.
- Perform scalability assessment: Record simulation runtimes as a function of graph size.

IV. RESULTS

In this section, we present the procedure for constructing the adjacency matrix of a second-order multilevel graph, defining its dynamic modeling through continuous functions and computing its structural derivative, as well as classifying fundamental relationships and analyzing their behavior both in the static matrix A and in its derivative \dot{A} . This methodological framework provides the quantitative foundation necessary to assess the evolution and simplification of complex systems, systematically integrating the visual approach proposed by Muñoz Garro [9] into a formal quantitative framework.

A. Static Adjacency Matrix

In this subsection, we rigorously develop the adjacency-matrix representation of a directed second-order multilevel graph. Let $G = (V, \psi_G)$, $|V| = n$ where ψ_G encodes all functional-dependency relations among vertices. The static adjacency matrix $A = [M_{ij}] \in \{0, 1\}^{n \times n}$ is defined entrywise by

$$M_{ij} = \begin{cases} 1, & \text{if there is a directed edge } i \rightarrow j, \\ 0, & \text{otherwise.} \end{cases}$$

Each “1” therefore represents a dependency from i to j . When G decomposes into k principal subgraphs of sizes n_1, \dots, n_k , we reorder vertices to reveal a $k \times k$ block structure:

$$A = \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \dots & A_{kk} \end{pmatrix},$$

Where each diagonal block $A_{ii} \in \{0, 1\}^{n_i \times n_i}$ captures intra-subgraph connectivity, whereas each off-diagonal block $A_{ij} \in \{0, 1\}^{n_i \times n_j}$ encodes inter-subgraph links.

Example: Let $V_1 = \{x, y, z\}$ and $V_2 = \{w, u\}$, with one internal edge $x \rightarrow y$ and one external edge $z \rightarrow w$. Ordering vertices (x, y, z, w, u) yields

$$A_{11} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, A_{12} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{pmatrix}, A_{21} = 0_{2 \times 3}, A_{22} = 0_{2 \times 2}$$

so that

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

B. Dynamic Modeling and Structural Derivative

This block-matrix construction serves as the basis for dynamic modeling. By treating A as a continuous function $A(t)$, we can compute its structural derivative

$$\dot{A} = \frac{dA(t)}{dt},$$

Which captures the instantaneous strengthening (positive entries) or weakening (negative entries) of dependencies. Furthermore, integrating $A(t)$ over any interval quantifies cumulative influence:

$$\int_{t_0}^{t_1} A(t) dt,$$

Where each I_{ij} measures the total or weighted duration of relation $i \rightarrow j$. In the following sections, we use this formalism both to classify fundamental relations and to compare behavior in the static matrix A versus its derivative \dot{A} , thereby providing a quantitative foundation for multilevel normalization and simplification.

C. Classification of Relations and Their Matrix Behavior

Point-to-Point

In a directed second-order graph, a point-to-point relation connects two simple vertices α and β in a one-to-one mapping:

$$\psi = \{\alpha \rightarrow \beta\} [9]$$

Its adjacency matrix, when ordering the vertices as $(\alpha \rightarrow \beta)$, is

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} [5].$$

If the weight of the edge $\alpha \rightarrow \beta$ varies continuously with parameter t as $w(t)$, the dynamic adjacency matrix and its structural derivative become

$$A(t) = \begin{pmatrix} 0 & w(t) \\ 0 & 0 \end{pmatrix}, \dot{A}(t) = \frac{dA(t)}{dt} = \begin{pmatrix} 0 & \dot{w}(t) \\ 0 & 0 \end{pmatrix} [10].$$

Given the elementary nature of this relation—and its foundational role in our taxonomy—we do not explore it further and instead proceed to relations of higher complexity.

Dynamic Relation (One-to-Many)

In a dynamic (one-to-many) relation, a composite vertex C connects to a set of simple vertices $\{s_1, \dots, s_m\}$. Formally, $\psi_C \ni \{C \rightarrow s_i; i = 1, \dots, m\}$.

When ordering the vertices as (C, s_1, \dots, s_m) , the adjacency matrix for this subgraph takes the block form

$$A = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

Where the first row has ones under each s_i and zeros elsewhere.

If each edge $C \rightarrow s_i$ carries a continuous weight $w_i(t)$, its structural derivative becomes

$$\dot{A}(t) = \begin{pmatrix} 0 & \dot{w}_1(t) & \dot{w}_2(t) & \dots & \dot{w}_m(t) \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

Where $\dot{w}_i(t) > 0$ indicates strengthening of $C \rightarrow s_i$ whereas $\dot{w}_i(t) < 0$ its weakening.

Interplanar Relations

Interplanar relations model many-to-many interactions between two composite nodes (or their subsystems), representing the highest complexity in our taxonomy [9]. We distinguish three variants:

Pure Interplanar

Let U and V be composite nodes with internal vertex sets

$$U_{\text{int}} = \{u_1, \dots, u_k\}, V_{\text{int}} = \{v_1, \dots, v_\ell\}.$$

The pure interplanar relation is

$$\psi_C = \{v_i \rightarrow v_j | i = 1, \dots, k; j = 1, \dots, \ell\}.$$

i.e., the Cartesian product $U_{\text{int}} \times V_{\text{int}}$. Ordering vertices as $(u_1, \dots, u_k, v_1, \dots, v_\ell)$, the adjacency matrix has the block form

$$A_{pure} = \begin{pmatrix} 0_{k \times k} & J_{k \times \ell} \\ 0_{\ell \times k} & 0_{\ell \times \ell} \end{pmatrix}$$

where $J_{k \times \ell}$ is the all-ones matrix.

Exponential Variant

This variant recurses into deeper levels: each u_i and v_i may themselves contain internal subsystems of sizes k_i and ℓ_i , respectively. The adjacency matrix A_{exp} then becomes a block matrix in which each block corresponding to (u_i, v_i) is itself a pure interplanar block of size $k_i \times \ell_i$, and diagonal blocks encode the internal structure of each subsystem [9]. Symbolically, one may write

$$A_{exp} = \begin{pmatrix} \ddots & \ddots & \vdots \\ \dots & 0_{k_i \times k_i} & 0_{k_i \times \ell_j} \\ \dots & 0_{\ell_j \times k_i} & 0_{\ell_j \times \ell_j} \end{pmatrix}$$

With pure-interplanar subblocks nested at each level.

Interplanar Dynamic

When a simple vertex s in one plane connects to every vertex in V_{int} , we obtain the dynamic interplanar variant:

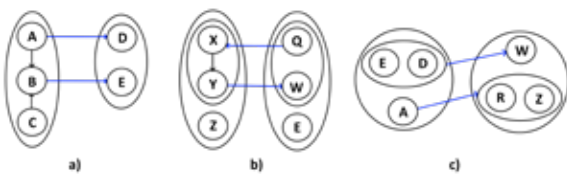
$$\psi_G \supseteq \{s \rightarrow v_j; j = 1, \dots, \ell\}$$

Ordering as (s, v_1, \dots, v_ℓ) , the static adjacency block is

$$A_{dyn} = \begin{pmatrix} 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}$$

For the interplanar dynamic variant—where every vertex in U_{int} connects to every vertex in V_{int} with time-dependent weights $w_{ij}(x)$ —the corresponding block matrix and its structural derivative are:

$$A(x) = \begin{pmatrix} 0 & w(x) \\ 0 & 0 \end{pmatrix}, A(x) = \frac{dA(x)}{dt} = \begin{pmatrix} 0 & \dot{w}(x) \\ 0 & 0 \end{pmatrix}$$



Interplanar Relations (a) Pure; (b) Exponential; (c) Dynamic.

Harmful Dependencies

In complex systems, certain relationships introduce noise and inefficiency by adding unnecessary complexity. We classify these harmful dependencies into four types—reciprocal, associative, redundant, and cyclic—and develop the first two categories below.

Reciprocal Dependencies

A reciprocal dependency between a simple vertex α and a composite vertex β retains both directions:

$$\{\alpha \rightarrow \beta, \beta \rightarrow \alpha\} \subseteq \psi_G$$

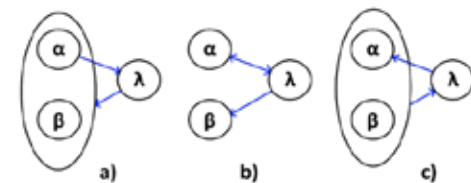
Ordering the vertices as (α, β) , the static adjacency matrix is

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

If the edge weights depend smoothly on a parameter t , via $w_{\alpha\beta}(t), w_{\beta\alpha}(t) \in C^1$, then

$$A(t) = \begin{pmatrix} 0 & w_{\alpha\beta}(t) \\ w_{\beta\alpha}(t) & 0 \end{pmatrix}, A(t) = \begin{pmatrix} 0 & \dot{w}_{\alpha\beta}(t) \\ \dot{w}_{\beta\alpha}(t) & 0 \end{pmatrix}$$

When the reverse flow $\beta \rightarrow \alpha$ is isolated from a compound dependency $(\alpha, \beta) \rightarrow \lambda$, the indistinguishable origin of λ requires that flow to be treated as a unidirectional (normal) dependency.



(a) Reciprocal Dependence; (b) Normal Dependence; (c) Normal Dependence.

Associative Dependencies

For

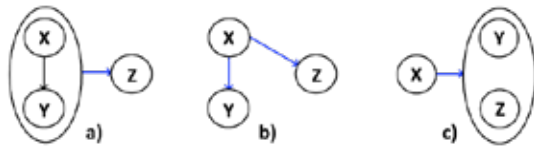
$$\{X \rightarrow Y, XY \rightarrow Z\} \subseteq \psi_G$$

Armstrong's union axiom [4] ensures $\{X \rightarrow Y, XY \rightarrow Z\} \equiv \{X \rightarrow YZ\}$, allowing us to replace ψ_G with $\psi'_G = \{X \rightarrow Y, X \rightarrow Z\}$ and eliminate $XY \rightarrow Z$ without loss provided at least one of $X \rightarrow Y$ or $X \rightarrow Z$ remains. Ordering vertices as (X, Y, Z) , the adjacency matrix becomes

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

If the weights of $X \rightarrow Y$ and $X \rightarrow Z$ vary with t as $f(t)$ and $g(t)$, then

$$A(t) = \begin{pmatrix} 0 & f(t) & g(t) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A(t) = \begin{pmatrix} 0 & f(t) & g(t) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$



(a) Associative Dependence; (b) Normal Dependence; (c) Normal Dependence (Dynamic).

Redundancies

Redundancies arise when a sequence of aligned dependencies creates unnecessary paths that, instead of adding new information, only increase graph complexity without improving structural integrity. In the coupling case, three (or more) edges arranged in series may occur, where the final edge is redundant because its existence can be inferred by transitivity from the first two [3]. Cyclical redundancies occur when dependencies form a closed loop—a cycle—that cannot be decomposed or simplified without losing essential information. In both scenarios, removing the superfluous edges normalizes the graph, reducing noise and inefficiency while preserving the validity of all functional dependencies.

Coupling Redundancy

In coupling redundancy, we consider $\psi_G = \{C \rightarrow A, A \rightarrow B, C \rightarrow B\}$. The edge $C \rightarrow B$ is redundant, since by transitivity of $C \rightarrow A$ and $A \rightarrow B$ its existence is already guaranteed, and so it only adds unnecessary complexity [3]. Ordering the vertices as (C, A, B) , the static adjacency matrix becomes

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

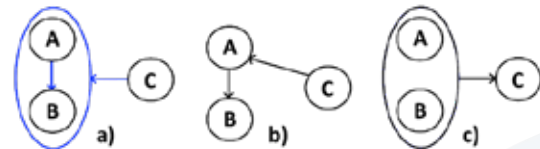
If each edge weight varies smoothly with parameter x via $w_{CA}(x), w_{AB}(x)$ and $w_{CB}(x) \in C^1$, then the time-varying adjacency matrix is

$$A(x) = \begin{pmatrix} 0 & w_{CA}(x) & w_{CB}(x) \\ 0 & 0 & w_{AB}(x) \\ 0 & 0 & 0 \end{pmatrix}$$

Where $w_{ij}(x) > 0$ indicates strengthening and $w_{ij}(x) < 0$ weakening of the edge $(i \rightarrow j)$ [11].

To normalize the graph, the graph is normalized by removing the $C \rightarrow B$ edge, yielding the reduced adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$



(a) Coupling Dependency; (b) Normal Dependency; (c) Normal Dependency.

Cyclical Dependencies

A cyclical dependency occurs when a sequence of functional relations forms a closed loop, and an additional direct edge is redundantly introduced into the cycle. For example, consider

$$\psi_G = \{X \rightarrow Z, Z \rightarrow B, B \rightarrow C, C \rightarrow Y, X \rightarrow Y\}$$

Where, by transitivity of $X \rightarrow Z \rightarrow B \rightarrow C \rightarrow Y$, the direct edge $X \rightarrow Y$ is redundant [4]. Ordering the vertices as (X, Z, B, C, Y) , the static adjacency matrix is

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Where the entry (1,5) marks the redundant arc $X \rightarrow Y$ [5]. Normalization simply removes this edge, preserving the path $X \rightarrow Z \rightarrow B \rightarrow C \rightarrow Y$ while eliminating the unnecessary cycle.

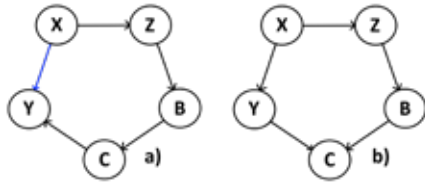
A false cycle resembles a closed loop but contains no removable edge without breaking connectivity. For instance:

$$\psi_G = \{X \rightarrow Z, Z \rightarrow B, B \rightarrow C, X \rightarrow Y, Y \rightarrow C\}$$

Here, no single edge can be removed without interrupting at least one path. With vertices ordered (X, Z, B, C, Y) , the adjacency matrix becomes

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

And since no “1” can be transitively deduced from the others, no normalization applies.



(a) True Cyclical Dependency; (b) Normal Dependency.

Numerical Validation via Simulations

A. Scenario 1: Synthetic Industrial Network

We construct a second-order digraph $G_1 = (V, \psi_G)$ consisting of five sequential subgraphs, each with ten vertices $\{n_{k,0}, n_{k,1}, \dots, n_{k,9}\}$ connected by edges $n_{k,i} \rightarrow n_{k,i+1}, i = 0, \dots, 8$, and additional skip-links $n_{k,9} \rightarrow n_{k,i+1}$ for $k = 0, \dots, 3$. This topology exemplifies pure point-to-point relations without redundancies or cycles. Our goal is to evaluate the behavior of the dynamic adjacency matrix $A(t)$, its derivative $\dot{A}(t)$, and its integral in a minimally complex setting [11].

Using a linear weight function $w(t) = 1 + 0.1t$ on every edge, we obtained:

Concordance (0%): All $w_{ij}(t) \cdot w_{ij}(t) \cdot w_{ij}(t)$ are identical, so classical and dynamic “top-10” node rankings do not significantly overlap.

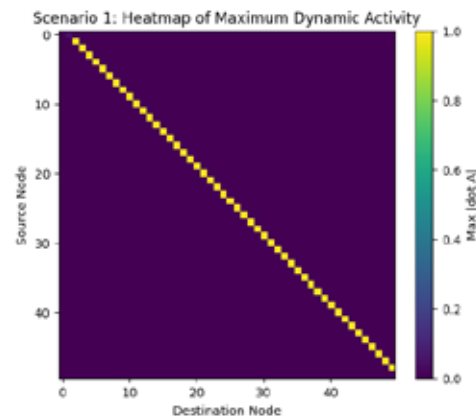
Concentration (1.6%): Only 1.6% of the matrix entries account for 80% of the integral weight, reflecting network sparsity.

Reduction (0%): No transitive edges exist to remove, consistent with a cycle-free, coupling-free topology.

Computation Time (~0.02 s): This demonstrates the method’s high efficiency for graphs up to 50 nodes.

Fig. 6 shows a heatmap of the maximum activity $\max_t |A_{ij}(t)|$. A single yellow band is observed on the superdiagonal—corresponding to the 49 sequential

edges—while all other entries remain zero.



Heatmap of Maximum Dynamic Activity for Scenario 1; yellow band indicates active super-diagonal edges; $\max_t |A_{ij}(t)| = 0.1$.

B. Scenario 2: Corporate Workflow Graph

We define a bipartite digraph $G_2 = (V, \psi_G)$ that models task assignment in an organization. The vertex set comprises four composite nodes (departments: HR, Finance, IT, Sales) and twenty simple nodes (tasks T_0, \dots, T_{19}). Each department connects randomly to five tasks via edges $D \rightarrow T$, exemplifying one-to-many dynamic relations [1]. We assess how the structural derivative $\dot{A}(t)$ identifies vertices of highest activity in a labor-assignment context [11].

Using a linear weight function $w(t) = 0.1t$ on every edge, we obtained:

Concordance (60%): 6 of the top-10 degree vertices (the four departments plus two highly connected tasks) coincide with the top-10 active vertices ranked by $\sum |A_{ij}|$.

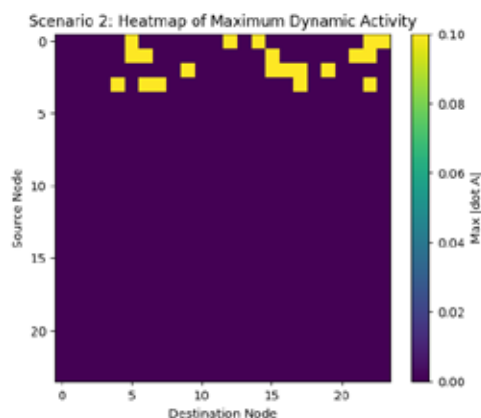
Concentration (2.95%): Only 2.95% of matrix entries account for 80% of the integral weight, indicating that a few high-flow links dominate the dynamics.

Reduction (0%): No transitively redundant edges were detected in this bipartite graph.

Computation Time (≈0.003 s): Confirms high efficiency even for organization-scale graphs.

Fig. 7 presents a heatmap of $\max_t |A_{ij}(t)|$. Bright rows correspond to the four

departments—origins of multiple outgoing edges—while task nodes, which do not redistribute flow, remain inactive (zero).



Heatmap of Maximum Dynamic Activity for Scenario 2; department rows highlighted; $\max_i |\dot{A}_{ij}(t)| = 0.1$.

Scenario 3: Synthetic Graph with Redundant Coupling

We construct a digraph $G_3 = (V, \psi_C)$ by combining an Erdős–Rényi random subgraph with a harmful coupling dependency. First, G_{ER} is generated on $n=50$ nodes with edge probability $p=0.05$. Then, a coupling subgraph on three nodes $\{C, A, B\}$ is appended, with edges $C \rightarrow A$, $A \rightarrow B$, and the transitively redundant $C \rightarrow B$ [1]. We assign a linear weight $w(t) = 1 + 0.1t$ to every edge and compute $\{A(t_k)\}$, $\dot{A}(t_k)$ and $\int A(t_k) dt$ numerically [11].

The resulting metrics are:

Concordance (40%): Four of the top-10 statically highest-degree vertices (random hubs and coupling nodes) coincide with the most active vertices by $\sum |A_{ij}|$.

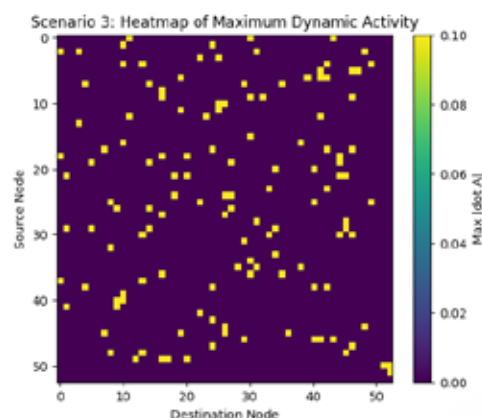
Concentration (3.74%): Only 3.74% of entries account for 80% of the total integral weight, indicating that a few critical edges drive most dynamics.

Reduction (53.44%): Over half of the original edges—including the redundant coupling edge and transitive edges in the random subgraph—are removed, validating the effectiveness of our matrix-based normalization.

Computation Time (≈ 0.012 s): Confirms high efficiency for medium-sized graphs.

Fig. 8 depicts a heatmap of $\max_t |\dot{A}_{ij}(t)|$. Dispersed yellow cells mark (i) edges

in the random subgraph—each with $\dot{w} = 0.1$ —and (ii) the two essential coupling edges $C \rightarrow A$ and $A \rightarrow B$ that remain after normalization.



Heatmap of Maximum Dynamic Activity for Scenario 3; yellow points indicate mixed random dynamics and the two preserved coupling edges; $\max_i |\dot{A}_{ij}(t)| = 0.1$.

V. CONCLUSIONS

We have demonstrated that traditional second-order graph normalization—historically qualitative and visually driven—can be significantly enhanced by reinterpreting it as a dynamic matrix framework. By assigning continuous weight functions to each edge, constructing the time-dependent adjacency matrix $A(t)$, and computing its structural derivative $\dot{A}(t)$ and integral $\int A(t) dt$, we provide a numerical scaffold that precisely measures both instantaneous change rates and accumulated influence over time, thereby extending beyond the static graph paradigm [1], [11].

Simulations across three contrasting scenarios—a linear industrial process network, a bipartite corporate-workflow graph, and a synthetic random graph augmented with a redundant coupling subgraph—validate the versatility of our approach. In Scenario 1, $\dot{A}(t)$ unambiguously highlights all 49 active links; in Scenario 2, it achieves 60% concordance with high-degree nodes and concentrates 3% of edges into 80% of total dynamic weight; and in Scenario 3, over 53% of redundant edges are removed, 40% overlap is observed between static and dynamic critical nodes, and only 3.7% of connections carry 80% of the integral weight.

These findings confirm that the structural derivative and matrix integral serve as complementary indicators to classical centrality metrics: they enable objective edge-prioritization for intervention

and quantify complexity reduction—up to 53% in the most demanding case—without compromising essential connectivity. Furthermore, simulation runtimes on the order of 0.01–0.2 s for graphs of up to 50 nodes demonstrate the method’s practicality for moderate-size applications and, with further algorithmic optimizations, its scalability to much larger networks.

Future work will explore nonlinear and stochastic weight functions to capture richer dynamics; dimensionality-reduction and intelligent-sampling techniques for networks with thousands of vertices; and integration into interactive visualization platforms for real-time parameter adjustment, thereby closing the loop between graphical intuition and numerical rigor. Together, these advances move the normalization of complex systems toward a unified visual, quantitative, and automatable framework.

REFERENCES

- [1] E. Muñoz Garro, “Normalization by second order graphs: a visual alternative to simplify systems,” *e-Ciencias de la Información*, vol. 11, no. 1, 2021. doi:10.15517/eci.v11i1.38790.
- [2] E. F. Codd, “A relational model of data for large shared data banks,” *Communications of the ACM*, vol. 13, no. 6, pp. 377–387, Jun. 1970. doi:10.1145/362384.362685.
- [3] W. W. Armstrong, “Dependency structures of database relationships,” in *Proceedings of Information Processing '74*. North-Holland, 1974, pp. 580–583.
- [4] W. W. Armstrong, Y. Nakamura, and P. Rudnicki, “Armstrong’s axioms,” *Journal of Formalized Mathematics*, vol. 14, no. 1, 2002.
- [5] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, vol. 290. London, U.K.: Macmillan, 1976.
- [6] B. Attallah, “Simplifying database normalization within a visual interactive simulation model,” *International Journal of Database Management Systems*, vol. 9, no. 3, pp. 57–69, 2017. doi:10.5121/ijdms.2017.9304.
- [7] G. Ausiello, A. D’Atri, and D. Saccà, “Graph algorithms for functional dependency manipulation,” *Journal of the ACM*, vol. 30, no. 4, pp. 752–766, 1983. doi:10.1145/2157.322404.
- [8] G. Di Battista, W. Didimo, M. Patrignani, and M. Pizzonia, “Drawing relational schemas,” in *Proceedings of Graph Drawing (GD)*.
- [9] D. Harel, “Statecharts: a visual formalism for complex systems,” *Science of Computer Programming*, vol. 8, no. 3, pp. 231–274, 1987. doi:10.1016/0167-6423(87)90035-9.
- [10] D. Kumar, A. Raj, and J. Dharanipragada, “Graphsteal: dynamic re-partitioning for efficient graph processing in heterogeneous clusters,” in *Proceedings of the 2017 IEEE 10th International Conference on Cloud Computing (CLOUD)*, 2017, pp. 439–446.
- [11] V. Pavlyshyn, “Dynamic block adjacency matrices in hierarchical graph models,” *Journal of Network Dynamics*, vol. 5, no. 2, pp. 123–140, 2025.
- [12] S. Shapiro, *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford, U.K.: Clarendon Press, 1991.
- [13] J. Ren, J. Schneider, M. Ovsjanikov, and P. Wonka, “Joint graph layouts for visualizing collections of segmented meshes,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 24, no. 9, pp. 2546–2558, 2017.
- [14] Q. Shi, G. Liu, K. Zheng, A. Liu, Z. Li, L. Zhao, *et al.*, “Multi-constrained top-K graph pattern matching in contextual social graphs,” in *Proceedings of the 2017 IEEE International Conference on Web Services (ICWS)*, 2017, pp. 588–595.



CONTROLES DE SEGURIDAD PARA LA PREVENCIÓN DE FUGA DE INFORMACIÓN (DLP) POR PARTE DE USUARIOS INTERNOS EN SISTEMAS HOSPITALARIOS DE LA CCSS EN COSTA RICA

I. Sáenz Córdoba

Escuela de Ciencias Exactas y Naturales, Universidad Estatal a Distancia, isaenzc@uned.ac.cr

ABSTRACT

Information security management in critical hospital infrastructures, such as the Costa Rican Social Security Fund (CCSS), faces unprecedented challenges following the cyberattacks suffered in 2022. This paper analyzes the implementation of Data Loss Prevention (DLP) controls aimed at mitigating risks originating from internal users, who, whether through negligence, the use of “Shadow IT,” or a lack of digital culture, represent the primary vector for sensitive data exposure. The study examines the need for a defense-in-depth strategy that integrates compliance with Law No. 8968 on the Protection of the Person Regarding the Processing of Personal Data with technical controls across data in use, in transit, and at rest. Finally, a hybrid model combining Identity and Access Management (IAM), activity monitoring, and administrative controls is proposed, concluding that technology alone is insufficient without a transformation in the organizational security culture.

Keywords: CCSS, DLP, Hospital Security, Insider Threat, Law 8968, Shadow IT, Data Protection, Cybersecurity.

RESUMEN

La gestión de la seguridad de la información en infraestructuras hospitalarias críticas costarricenses, como la de la Caja Costarricense de Seguro Social (CCSS), enfrenta desafíos sin precedentes tras los ciberataques sufridos en 2022. Por lo que este artículo analiza la implementación de controles de prevención de fuga de datos (DLP) orientados a mitigar los riesgos provenientes de usuarios internos; quienes, ya sea por negligencia, uso de Shadow IT o falta de cultura digital, representan el principal vector de exposición de datos sensibles. Además, se examina la necesidad de una estrategia de defensa en profundidad que integre el cumplimiento de la Ley N.º 8968 de Protección de la Persona frente al Tratamiento de sus Datos Personales con controles técnicos en los estados de uso, tránsito y reposo de la información. Finalmente, se propone un modelo híbrido que combina la segmentación de identidades (IAM), el monitoreo de actividad y controles administrativos, y se concluye que la tecnología por sí sola es insuficiente sin una transformación en la cultura organizacional de seguridad.

Palabras clave: CCSS, DLP, seguridad hospitalaria, amenaza interna, Ley 8968, Shadow IT, protección de datos, ciberseguridad.

INTRODUCCIÓN

El uso profesional y ético de las herramientas tecnológicas y sistemas informáticos facilitados por la organización en el entorno hospitalario implica que tanto el usuario directo (médicos, enfermeras, administrativos) como el usuario indirecto (pacientes del sistema de salud público costarricense) representen el mayor vector de riesgo para la fuga de información; no necesariamente por malicia, sino por negligencia o Shadow IT (uso de herramientas no autorizadas para agilizar el trabajo, tales como smartphones, tablets o equipos tecnológicos personales), o bien por hacer uso de redes ajenas facilitadas por la institución.

Estas acciones, si bien es cierto son parte de la cotidianidad del personal médico y administrativo, aunque parezcan bien intencionadas, pueden derivar en la fuga accidental de información sensible, especialmente en sistemas donde convergen datos clínicos, diagnósticos, enfermedades, historiales y resultados médicos altamente sensibles de los pacientes de la CCSS. Por lo cual, surge la importancia de la implementación de controles DLP adaptados al comportamiento interno con los mejores estándares de seguridad informática.

Tras el ciberataque sufrido en Costa Rica en mayo del 2022 al seguro social del país se deduce que la CCSS ha incrementado sus medidas de ciberseguridad y hoy por hoy concentra sus esfuerzos en el control de los usuarios internos, ya que este aspecto resulta clave para reducir el riesgo de fuga de información tanto de colaboradores como de los usuarios de la Caja, sin embargo el crear una cultura de seguridad que combine tecnología y responsabilidad individual para proteger la información crítica de la institución tiene muchos aspectos por mejorar.

En los pasillos hospitalarios, áreas de salud, sucursales y otras oficinas adyacentes, coexisten decenas de colaboradores, desde médicos y enfermeras hasta personal administrativo; quienes, de alguna u otra manera, cuentan con un perfil tecnológico único. Es un entorno donde interactúan diariamente usuarios capaces de gestionar complejos sistemas clínicos con información altamente sensible de pacientes e inclusive otros usuarios del seguro social costarricense que aún enfrentan barreras básicas en la interacción digital, falta de recursos, capacitaciones, problemas en la red y otros factores adversos que permiten el filtrado de los datos.

Esta profunda brecha de habilidades transforma al personal interno en un elemento crítico de la ecuación de riesgo y la falta de conciencia digital sobre las graves consecuencias de una eventual fuga de información involuntaria y desconocimiento de la Ley 8968, Ley de Protección de la Persona frente al Tratamiento de sus Datos Personales.

En perspectiva, la Auditoría Interna de la CCSS indica que aún existen cuentas con asignación excesiva de roles, duplicidad de funciones y perfiles sin descripción, lo que refleja un manejo ineficiente de privilegios dentro de varios de los sistemas institucionales a lo largo y ancho del país [1].

Por su parte, la implementación de controles DLP repercute de manera confrontativa en infraestructuras críticas de la institución que, debido a su antigüedad, representan desafíos significativos vinculados a la infraestructura tecnológica y a la dinámica operativa del personal. Así mismo, impide la adopción de agentes modernos de DLP, por lo que la estrategia debe apoyarse en controles perimetrales de red como medida complementaria para incrementar los niveles de seguridad informática en los nosocomios y algunas áreas de salud.

La resistencia al cambio y la fricción operativa constituyen barreras críticas, exacerbadas por la alta presión y los esquemas de rotación continua (24/7) del personal médico. Estas condiciones laborales propician la adopción de Shadow IT como mecanismo para agilizar procesos; por consiguiente, la implementación de controles de seguridad debe ser gradual, priorizando fases de monitoreo y alerta antes de habilitar bloqueos restrictivos.

Adicionalmente, las limitaciones de conectividad en zonas rurales exigen el despliegue de agentes de endpoint ligeros y autónomos (offline). La inestabilidad de la red en estas áreas fomenta prácticas de alto riesgo, como el uso compartido de cuentas genéricas y el traslado de datos sensibles de pacientes mediante dispositivos de almacenamiento extraíble (USB), lo cual compromete la integridad y confidencialidad de la información.

La Caja Costarricense de Seguro Social (CCSS) presenta una infraestructura tecnológica caracterizada por sistemas críticos obsoletos, equipos con actualizaciones y parches atrasados, así como una carencia de segmentación robusta en sus redes. Además, se evidencian limitaciones en los procesos de continuidad de negocio y una calidad del servicio informático deficiente en determinadas regiones del país; factores que incrementan significativamente la superficie de ataque.

Tras el incidente de ransomware dirigido inicialmente al Ministerio de Hacienda de Costa Rica y, después, a la CCSS, el grupo Hive logró comprometer los sistemas institucionales, cifrar información sensible, paralizar servicios esenciales y generar un impacto de alcance nacional [2]. En respuesta, la CCSS activó planes de contingencia, inició un proceso gradual de restauración de sistemas y fortaleció sus capacidades de ciberseguridad. Este fortalecimiento se vio potenciado por la intervención de los Estados Unidos con un presupuesto inicial de \$25 millones y la implementación de políticas públicas orientadas a mejorar la postura de seguridad nacional, aunque con resultados limitados en términos de resiliencia tecnológica y primeros pasos en ciberseguridad [3]

I. CONTEXTO NORMATIVO BASADO EN LA NORMA DE COSTA RICA

La implementación de controles de seguridad orientados a prevenir la fuga de información en los sistemas hospitalarios de la Caja Costarricense de Seguro Social (CCSS) debe fundamentarse en el marco normativo vigente en Costa Rica y en la clasificación institucional de datos críticos. Dicho marco establece los criterios para determinar qué información requiere protección, bajo qué principios regulatorios y con qué responsabilidades legales y operativas. En este contexto, entidades nacionales como la Agencia de Protección de Datos de los Habitantes (PRODHAB), el Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT), entre otras organizaciones, constituyen referentes normativos y técnicos para el tratamiento de la información personal y la gestión de la seguridad de los datos.

La Ley N.º 8968 constituye el marco normativo fundamental en Costa Rica en materia de privacidad y protección de datos personales. Esta legislación clasifica explícitamente la información clínica, expedientes médicos, diagnósticos y datos biomédicos como datos sensibles; los cuales requieren la implementación de salvaguardas técnicas, administrativas y organizativas de nivel superior, con el fin de limitar su divulgación.

Asimismo, la Ley 8968 se reconoce como un referente pionero en la regulación costarricense sobre el tratamiento y la usabilidad de la información personal. Un principio central de esta normativa es el consentimiento informado, establecido como requisito para el tratamiento o transferencia de datos sensibles, salvo en casos de excepción definidos por la misma ley.

La normativa también estipula que la divulgación indebida o el manejo inadecuado de dichos datos acarrea sanciones administrativas y económicas, aplicables tanto a las instituciones como a los colaboradores responsables. Dicho consentimiento debe ser formalizado mediante la firma y aprobación del usuario final, garantizando así la legitimidad del uso de su información privada [4].

Complementariamente al marco legal principal, la gestión de la información en el sector salud se rige por lineamientos de la PRODHAB que exigen controles reforzados para evitar fugas o accesos no autorizados, dependiente de la mala manipulación de los datos y, en caso de que se demuestre inconsistencias, las sanciones varían según la gravedad de la falta; además, pueden incluir multas basados en salarios bases, eliminación de datos y, en otros casos más estrictos, se podría incluir órdenes administrativas y sanciones más graves que van de 5 a 20 salarios base.

Por otra parte, la normativa interna de la Caja Costarricense de Seguro Social (CCSS), la cual regula estrictamente la seguridad y el control de acceso a sistemas institucionales críticos, como EDUS, ARCA y SICERE [5], [6], procura alinearse con los pilares de la ciberseguridad y, sobre todo, con la confidencialidad obligatoria estipulada en la Ley General de Salud y la Ley de Delitos Informáticos (Ley N.º 9048); la cual sanciona el acceso indebido, la sustracción, alteración o destrucción de información protegida [7], [8].

II. ESTRATEGIA DE CONTROLES TÉCNICOS DE DLP

Para mitigar el riesgo asociado a usuarios internos, ya sea por negligencia, malicia, debilidad o pésimos controles en los sistemas informáticos, es imperativo implementar controles de prevención de pérdida de datos (DLP) que abarquen los tres estados fundamentales de la información. En primer lugar de esta lista, se mencionan por su origen en inglés endpoint (datos en uso), donde las medidas deben centrarse en la restricción de periféricos no autorizados, el control del portapapeles y la implementación de impresión segura con autenticación. En segundo lugar, se menciona que la protección de datos en movimiento requiere la inspección profunda de tráfico cifrado (SSL/TLS), que hace posible el bloqueo de almacenamiento en nubes personales externas y el filtrado de correo saliente basado en contenido sensible.

Por último, para salvaguardar los datos en reposo, es crucial el despliegue de herramientas de descubrimiento y clasificación automatizada en servidores de archivos, junto con el monitoreo de actividad en bases de datos (DAM) para alertar sobre consultas masivas o comportamientos anómalos de usuarios con altos privilegios [9]. Para poner en perspectiva lo indicado anteriormente, se resumen los tres estados del dato (Figura 1).



Figura 1. Los tres estados del dato: en reposo, en tránsito y en uso [10].

III. CONTROLES DE IDENTIDAD Y ACCESO (IAM)

Es importante mencionar que, en el mayor de los casos, la incidencia de fugas de información suele correlacionarse directamente con la asignación de privilegios excesivos a los usuarios de las plataformas tecnológicas; lo que hace ineludible adoptar prácticas de Principio de Mínimo Privilegio (PoLP) y Control de Acceso Basado en Roles (RBAC). El Principio del Mínimo Privilegio (PoLP) es el mecanismo de seguridad que establece que cualquier entidad (como un usuario, aplicación o sistema) debe operar con la menor cantidad de permisos necesarios para realizar su tarea específica; es decir, el médico del EB AIS u hospital deberá únicamente visualizar la información del área de salud competente a la que están adscritos los pacientes, o bien a nivel general, si se trata del ámbito hospitalario. Por otra parte, el Control de Acceso Basado en Roles (RBAC) implementa un mecanismo o modelo de autorización en el que únicamente se deben agrupar los permisos en roles definidos (por ejemplo, “Lector,” “Escritor,” “Administrador”), es decir, si es un médico administrativo o un operativo, claramente debería contar con los roles de usuarios según sus responsabilidades laborales y no “mezclar” en qué se debería tener acceso y qué no.

En entornos hospitalarios complejos como el de la CCSS, esta segmentación dentro de plataformas como el EDUS debe ser granular, garantizando que el personal administrativo (mayormente departamentos de registros de estadística) acceda únicamente a datos logísticos y de agenda; mientras que perfiles clínicos, como farmacéuticos, visualicen exclusivamente prescripciones sin acceso a diagnósticos psiquiátricos o quirúrgicos detallados. Para blindar estas restricciones lógicas, es necesario habilitar mecanismos de Autenticación Multifactor (MFA) en todos los módulos críticos de la red interna y mitigar las vulnerabilidades de seguridad física mediante políticas de gestión de sesiones que fuercen el cierre automático (time-out) tras periodos breves de inactividad, impidiendo así

el acceso oportunista a terminales desatendidas en áreas de alta rotación como emergencias, farmacia, laboratorio y rayos x; además, que pueda darse la posibilidad de que un tercero tenga acceso a información clasificada sin el permiso adecuado de manera intencional o accidental [9], [11].

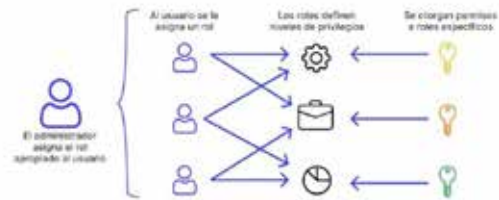


Figura 2. El control de acceso basado en roles [12].

IV. CONTROLES ADMINISTRATIVOS Y PROCESOS INTERNOS

La eficacia de la infraestructura tecnológica de seguridad se ve severamente comprometida, si no se acompaña de una cultura organizacional robusta, dado que el factor humano persiste como el vector de ataque más crítico en la cadena de eliminación de seguridad cibernética Cyber Kill Chain en inglés. Por consiguiente, es necesario establecer controles administrativos más robustos de los existentes que trasciendan la tecnología, comenzando por la formalización de Acuerdos de Confidencialidad (NDA) vinculantes que integren cláusulas explícitas sobre las responsabilidades derivadas de la Ley N.º 8968, aplicables no solo al personal de planta regular, sino extendiéndose obligatoriamente a practicantes o pasantes, proveedores y contratistas externos que, de alguna u otra forma, cuentan con acceso a las plataformas institucionales por la naturaleza de sus puestos de trabajo.

Simultáneamente, para mitigar la susceptibilidad ante ataques de ingeniería social, se deben implementar programas de concienciación continuos y simulacros de phishing que permitan evaluar y corregir el comportamiento de los usuarios frente a amenazas dirigidas. Finalmente, en el plano físico, la seguridad de la información debe reforzarse mediante una política estricta de “escritorio limpio”, erradicando prácticas de riesgo como la exposición de contraseñas en notas adhesivas o el abandono de expedientes clínicos en áreas de libre tránsito como las estaciones de enfermería, previniendo así la fuga de datos por observación directa o descuido que, lamentablemente, por experiencia del autor como paciente en distintos servicios de salud, se observan siempre en las impresoras direcciones IP y, en el peor de los casos, se observan pequeños papeles en blanco con anotaciones a mano como usuarios o claves a los distintos sistemas y a la red [13], [14].



Figura 3. Tomada de Facebook CCSS con posibles usuarios y claves debajo del monitor [15].

V. REFLEXIÓN FINAL

Se concluye que una estrategia efectiva de prevención de fuga de información (Data Loss Prevention, DLP) no puede limitarse exclusivamente a la implementación de controles tecnológicos. Por el contrario, requiere un enfoque integral que incorpore auditorías de seguridad física y programas de concienciación dirigidos a los usuarios. Este enfoque holístico permite mitigar el factor humano, identificado como el vector de riesgo más crítico y recurrente en la protección de datos sensibles. En este sentido, los usuarios finales constituyen el eslabón más vulnerable de la cadena de seguridad, dado que, de manera consciente o inconsciente, pueden facilitar el acceso no autorizado a plataformas tecnológicas que gestionan información altamente sensible. [16]

REFERENCIAS

A. Segura, «CCSS gasta más de \$1.4 millones en licencias de sistema informático que no se utilizan», CR Hoy. Accedido: 21 de enero de 2026. [En línea]. Disponible en: <https://crhoy.com/nacionales/ccss-gasta-mas-de-1-4-millones-en-licencias-de-sistema-informatico-que-no-se-utilizan/>

Caja Costarricense de Seguro Social (CCSS), «Oficio de Advertencia sobre la exposición reciente a ataques cibernéticos a la CCSS.» Costa Rica, 31 de mayo de 2022. [En línea]. Disponible en: <https://www.ccss.sa.cr/arc/auditoria/informes/AD-ATIC-067-2022.pdf>

T. Gómez, «EE.UU. y Costa Rica firman plan de \$25 millones para ciberseguridad», El Observador CR. Accedido: 21 de enero de 2026. [En línea]. Disponible en: <https://observador.cr/ee-uu-y-costarica-firman-plan-de-25-millones-para-ciberseguridad/>

Asamblea Legislativa de Costa Rica, «Ley 8968: Ley de Protección de la Persona frente al tratamiento de sus datos personales [Diario Oficial La Gaceta N.º 170]», Sistema Costarricense de Información Jurídica. Accedido: 21 de enero de 2026. [En línea]. Disponible en: https://pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_text_o_completo.aspx?param1=NRTC&nValor1=1&nValor2=70975&nValor3=85989&strTipM=TC

Agencia de Protección de Datos de los Habitantes (PRODHAB), «Guías y directrices sobre medidas de seguridad para el tratamiento de datos personales». San José, Costa Rica.

Caja Costarricense de Seguro Social (CCSS), «Normativa de Seguridad de la Información y Control de Acceso a Sistemas Institucionales». San José, Costa Rica.

Asamblea Legislativa de Costa Rica, «Ley 5395: Ley General de Salud», Sistema Costarricense de Información Jurídica. Accedido: 21 de enero de 2026. [En línea]. Disponible en: https://pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_text_o_completo.aspx?nValor1=1&nValor2=6581

Asamblea Legislativa de Costa Rica, «Ley 9048: Reforma de la Sección VIII, Delitos Informáticos y Conexos, del Título VII del Código Penal», Sistema Costarricense de Información Jurídica. Accedido: 21 de enero de 2026. [En línea]. Disponible en: https://pgrweb.go.cr/scij/Busqueda/Normativa/Normas/nrm_text_o_completo.aspx?nValor1=1&nValor2=73583

National Institute of Standards and Technology, «Security and Privacy Controls for Information Systems and Organizations», NIST, Gaithersburg, MD; U.S. Department of Commerce, NIST Special Publication (SP) 800-53 Rev. 5, dic. 2020. doi: 10.6028/NIST.SP.800-53r5.

sealpath, «Protegiendo la información en sus tres estados», Sealpath. Accedido: 21 de enero de 2026. [En línea]. Disponible en: https://www.sealpath.com/es/blog/tres_estados_info/

D. F. Ferraiolo y D. R. Kuhn, «Role-Based Access Controls», presentado en 15th National Computer Security Conference, Baltimore, USA, 1992, pp. 554-563. [En línea]. Disponible en: <https://csrc.nist.gov/files/pubs/conference/1992/10/13/rolebased-access-controls/final/docs/ferraiolo-kuhn-92.pdf>

I. Novikov, «¿Qué es RBAC (control de acceso basado en roles)?», Wallarm. Accedido: 21 de enero de 2026. [En línea]. Disponible en: <https://lab.wallarm.com/what/que-es-rbac-control-de-acceso-basado-en-roles/?lang=es>

International Organization for Standardization (ISO), «ISO/IEC 27002:2022 - Information security, cybersecurity and privacy protection — Information security controls». Ginebra, Suiza. Accedido: 21 de enero de 2026. [En línea]. Disponible en: <https://www.iso.org/standard/75652.html>

EL PARADIGMA DE LA INGENIERÍA DE CALIDAD DE SOFTWARE ASISTIDA POR IA GENERATIVA: EFICIENCIA, AUTOMATIZACIÓN Y TRANSFORMACIÓN SECTORIAL

Ing. Kevin Fernandez

QA mid & Black belt six sigma
 Kevin.fernandeze@cpic.cr
 Cartago, COSTA RICA

RESUMEN

La transición de la programación tradicional a la programación asistida por Inteligencia Artificial Generativa (GenAI) ha forzado una evolución paralela en el aseguramiento de calidad (QA). Este artículo analiza el impacto de la IA en la validación de software, comparando los métodos manuales frente a los asistidos por algoritmos, y explora aplicaciones críticas en los sectores de banca, finanzas y comercio. Finalmente, se abordan los límites éticos y el alcance real de estas herramientas en la mejora de procesos.

ABSTRACT

The transition from traditional programming to Generative Artificial Intelligence (GenAI) assisted programming has forced a parallel evolution in Software Quality Assurance (QA). This article analyzes the impact of AI on software validation, comparing manual methods against algorithm-assisted approaches, and explores critical applications within the banking, finance, and retail sectors. Finally, it addresses the ethical boundaries and the actual scope of these tools in process improvement.

Palabras clave: Inteligencia Artificial Generativa (GenAI), Aseguramiento de Calidad (QA), automatización cognitiva, SDLC, FinTech, pruebas de software, auto-curación, ética de la IA, CI/CD.

Keywords: Generative AI, Quality Assurance (QA), Cognitive Automation, SDLC, FinTech, Software Testing, Self-healing, AI Ethics, CI/CD.

I. INTRODUCCIÓN: DE LA CODIFICACIÓN A LA VALIDACIÓN

Como se expone en la literatura reciente sobre programación asistida [1], el uso de herramientas como GitHub Copilot o ChatGPT no solo acelera la escritura de código, sino que redefine cómo entendemos la lógica del programa. En la ingeniería de calidad, esto se traduce en una capacidad sin precedentes para generar casos de prueba, scripts de automatización y detección temprana de vulnerabilidades. La IA ya no es solo un oráculo de consulta, sino un colaborador activo en el ciclo de vida de desarrollo de software (SDLC).



Fig. 1. El ciclo de vida del defecto con Inteligencia Artificial Generativa (GenAI): Optimización del SDLC con Inteligencia Artificial (IA)

II. IMPACTO: CALIDAD MANUAL VS CALIDAD ASISTIDA POR IA

La diferencia entre el enfoque tradicional y el asistido por IA marca un punto de inflexión en el Time-to-Market.

A. QA Manual: El Factor humano y sus límites

El QA manual es de suma importancia para la experiencia de usuario (UX), pero presenta deficiencias en:

Escalabilidad: Imposible de ejecutar en entornos de integración continua (CI/CD) de alta frecuencia.

Error humano: Fatiga en pruebas de regresión repetitivas.

Costo: Alto consumo de horas-hombre para tareas de bajo valor intelectual.

Tabla 1. Tabla comparativa: QA Manual vs. QA Asistido por IA

Característica	QA Manual Tradicional	QA Asistido por GenAI
Velocidad de ejecución	Limitada por la jornada humana	Síncrona con pipelines de CI/CD
Mantenimiento de scripts	Manual y propenso a errores	Self-healing (Auto-curación)
Datos de prueba	Datos estáticos o limitados	Generación sintética masiva
Cobertura de seguridad	Muestreo periódico	Análisis continuo de vulnerabilidades
Costo a largo plazo	Escala lineal con el personal	Escala marginal tras la implementación

B. QA Asistido por IA: La automatización cognitiva

A diferencia de la automatización basada en reglas (scripts rígidos), la IA ofrece:

Pruebas de auto-curación (self-healing): Si un elemento de la interfaz cambia (ej. un botón cambia de ID), la IA identifica el cambio y ajusta el script automáticamente, reduciendo el mantenimiento en un 70 %.

Generación sintética de datos: Creación de millones de registros de prueba que imitan el comportamiento real sin comprometer datos sensibles.

Detección de anomalías: Identificación de patrones de error que pasarían desapercibidos para el ojo humano en grandes volúmenes de logs.



Fig. 2. Impacto en el Time-to-Market antes y después de la autocuración de scripts.

III. APLICACIONES REALES: SECTORES ESTRATÉGICOS

A. Sector banca y finanzas

En un entorno donde el error cuesta millones y la seguridad es innegociable, la IA actúa como un filtro de alta precisión.

Pruebas de seguridad: La IA analiza el código en busca de patrones de inyección SQL o vulnerabilidades de día cero antes del despliegue.

Validación de transacciones: Simulación de picos de carga masivos para probar la resiliencia de los núcleos bancarios ante eventos como el “Viernes Negro” o cierres fiscales.

Cumplimiento regulatorio: Verificación automatizada de que cada flujo de pago cumple con normativas internacionales (ISO 20022).

B. Sector comercial y retail

La calidad aquí se mide en la fluidez del “Customer Journey”.

Pruebas de regresión multicanal: La IA valida que una promoción se aplique correctamente de forma simultánea en la web, aplicación móvil y puntos de venta físicos.

Pruebas A/B automatizadas: Algoritmos que prueban miles de variaciones de una interfaz para determinar cuál genera menos errores de conversión y mejor respuesta del sistema.



Fig. 3. Aplicaciones en sectores estratégicos banca y retail

IV. MEJORA DE PROCESOS MEDIANTE IA GENERATIVA

Siguiendo los principios de optimización tecnológica, la integración de GenAI en las pruebas permite:

Refactorización de pruebas legado: Convertir scripts antiguos en lenguajes modernos de forma automatizada.

Documentación de calidad: Generación automática de reportes de errores técnicos explicados en lenguaje natural para los stakeholders financieros, lo que facilita la toma de decisiones basada en riesgos.

V. CONCLUSIONES

Alcance de la IA: La IA no reemplaza al ingeniero de calidad; lo eleva a un rol de arquitecto de pruebas. El alcance actual permite la automatización de la ejecución y el diseño básico, pero el pensamiento estratégico y la validación ética siguen siendo prerrogativas humanas.

Desafíos éticos: Si la IA se entrena con código defectuoso o sesgado, las pruebas que genere heredarán esos errores, creando una falsa sensación de seguridad.

Propiedad intelectual: Como se discute en [1], el código generado puede contener fragmentos sujetos a licencias. El QA debe validar que las herramientas de IA no introduzcan riesgos legales en el software financiero.

Transparencia: En finanzas, “la IA dijo que estaba bien” no es una defensa válida ante un regulador. La IA explicable es el siguiente paso necesario para auditar cómo se validó el software.

REFERENCIAS

[1] D. Zingaro y L. Porter, Learn AI-Assisted Python Programming, 2.^a ed., Shelter Island, NY: Manning Publications, 2024.

[2] R. S. Pressman y B. R. Maxim, Software Engineering: A Practitioner’s Approach, 9.^a ed., New York, NY: McGraw-Hill Education, 2020.

[3] J. G. Enríquez y J. J. Gutiérrez, «Quality Assurance en la era de la Transformación Digital», Journal of Software Engineering Research and Development, vol. 12, n.º 4, pp. 45-60, 2024.

[4] International Software Testing Qualifications Board (ISTQB), Certified Tester AI Testing (CT-AI) Syllabus, Bruselas: ISTQB, 2023.

[5] S. Russell y P. Norvig, Artificial Intelligence: A Modern Approach, 4.^a ed., Hoboken, NJ: Pearson, 2021.

[6] M. Fowler, Refactoring: Improving the Design of Existing Code, 2.^a ed., Boston, MA: Addison-Wesley, 2019.



REVIEW AND ORGANIZATIONAL ANALYSIS OF THE IT DEPARTMENT AT THE NATIONAL UNIVERSITY OF COSTA RICA

Mag. Steven Cruz-Sancho

*Computer Engineer, Academic Staff, Sarapiquí Campus, UNA
STEVEN.CRUZ.SANCHO@una.cr*

Mag. Angie Gómez-Mora

*Administrator, Academic Staff, Coto Campus, UNA
ANGIE.GOMEZ.MORA@una.cr*

Mag. Olivier Blanco-Sandí

*Computer Engineer, Academic Staff, Sarapiquí Campus, UNA
OLIVIER.BLANCO.SANDI@una.cr*

Mag. Rebeca Elizondo-Pizarro

*Administrator, Academic Staff, Sarapiquí Campus, UNA
MARIA.ELIZONDO.PIZARRO@una.cr*

ABSTRACT

This article presents the results of a research project conducted to perform a situational diagnosis of the Information Technology department at the National University of Costa Rica, Sarapiquí Campus. The objective of the diagnosis is to identify and analyze the Information Technology department in relation to its objectives, goals, occupational structure, and strategies, particularly regarding its support for the university's strategic planning and human resources. Methodologically, a descriptive research design with a qualitative approach was employed. From the perspective of data collection and results analysis, a qualitative emphasis was adopted, involving the selection of participants linked to the IT department. These individuals were provided with a set of self-designed instruments (questionnaires and interviews) to gather and analyze information. Among the main findings, it was identified that the IT department supports the essential functions and services of the Campus; however, it lacks formal planning guidelines and adherence to established standards in the field. Additionally, although the department makes significant efforts in carrying out its tasks in alignment with its occupational profile to support the institution, it is necessary to document these actions through information systems.

Keywords: Information technologies, human resources, strategic planning, organizational analysis, information systems.

I. INTRODUCTION

Every organization, in pursuing the achievement of its goals, aims to operate efficiently so that it can develop its products or services in a harmonious manner. To achieve this, it is important to understand the current situation in which its departments operate and to validate their daily operations based on the available human resources and the established job manual within the organizational structure.

Organizational effectiveness can be conceptualized as the degree to which an organization achieves its goals [1][2]. From another perspective,

organizational effectiveness is a challenge that requires adopting a broad view of performance and analyzing important strategic results, which in turn requires examining functions at both the individual and departmental levels [3].

Beyond designing its organizational structure and functional divisions at the human resources level, every organization must seek the continuous improvement of each of its departments. Therefore, it is imperative to conduct regular and periodic reviews of both service delivery efficiency and process structures.

Authors in a systematic literature review study point out that there should be a close relationship between technology and human resources practices, emphasizing their interconnection and identifying that both reinforce each other in generating benefits in the current technological era. This demonstrates the necessity of analyzing and validating employees' roles and the value they contribute to the organization from a human resources perspective [4].

In this context, focusing on the organizational setting of the National University of Costa Rica, specifically within its Huetar Norte and Caribbean Regional Section, Sarapiquí Campus, this research presents the identification and analysis of the current situation of the information technology department in terms of its operations and organizational structure.

II. THEORETICAL FRAMEWORK

The situational review of an information technology department constitutes one of the phases in the development of a strategic information technology plan, as it allows for an understanding of the functioning of the processes and services provided within the department.



Fig. 1. Methodology for Strategic Planning

Fig. 1 illustrates the adaptation of Anita Cassidy's methodological proposal, which provides guidance on how to develop a strategic information technology plan within an organization. The focus of the present study is placed on Phase 1, specifically on the identification and analysis of the current situation of the technology department.

For the implementation of a situational assessment within an information technology department, it is essential to understand the context in which it operates, its human resources organization, and its organizational structure, as well as the services and initiatives it offers, since these provide the necessary basis for departmental analysis [5].

In this way, it becomes possible to develop alternative recommendations and solutions.

Below, the main theoretical references related to topics such as strategic planning, information technology departments, and information regarding the National University of Costa Rica are discussed. These references help to understand the relevance of the department and its role in carrying out its functions.

2.1 Context of Strategic Planning in Information Technology

Given the absorbing and dynamic labor market in which organizations must compete, evaluating their strategies and departmental functioning should be a priority and a constant task. For this reason, to maintain competitiveness, constant measurement through alternatives such as control indicators and feedback is necessary.

To understand the concept of strategic planning and how it supports the work carried out by the information technology department, it is first important to examine the particularities of its components. Therefore, the concepts of information technology and strategic planning will be analyzed in the following sections.

2.2 Information Technology

Information and communication technologies are understood as computational and information tools used to process, store, synthesize, and retrieve information in various forms and, as a result, can be considered means that support digitalization [6]. It is determined that:

“As information and communication technologies, it can be mentioned all the technological resources used to manage, process, store, and present information, which have been transformed into tools that automate processes previously performed by humans, and this is reflected in the role they play in mediating teaching and learning processes” [7, p. 4].

Based on these definitions, it is determined that information technologies are tools and equipment capable of storing, analyzing, and presenting information with the purpose of automating processes within the organization.

2.3 Strategic Planning

To understand what “strategic planning” means, it is important to understand what each component involves; therefore, it is necessary to be clear about the concepts of planning and strategy.

Planning can be seen as the process of establishing a clear direction and creating an appropriate environment for managing a business, institution, or social organization in an informed and innovative manner [8]. In other words, planning involves defining an organization’s goals, selecting the most appropriate strategies, and organizing people and tasks so that all elements work together effectively. Essentially, it consists of determining what needs to be done and how to accomplish it. [9].

In one of the historically accepted definitions, strategy can be considered the creation of a unique and highly valued position that integrates a set of activities, with the particularity that it must differ from those of other organizations in order to generate added value and provide a distinctive identity through its products or services compared to competitors [10]. From another perspective, strategy is the determination of a company’s fundamental purpose, its lines of action, and the allocation of resources to achieve established goals. Finally, it is necessary to address the importance of strategic management through the involvement of human resources and the recognition of their role in carrying out their functions, since their performance contributes to the achievement of the objectives established by the organization [12].

Having clarified the concepts of planning and strategy, strategic planning can be defined as the harmonious coordination of a set of actions aimed at positively contributing to an organization’s goals.

2.4 Strategic Planning of Information Technology

The information technology department is a unit that functionally supports the organization; for this reason, its efforts should be directed toward organizational benefits, bearing in mind that technology is a means rather than an end. Therefore, the strategic planning of the IT department must be closely linked to the achievement of the organization’s goals and objectives, in this case, those of the Sarapiquí Campus of the National University of Costa Rica.

Strategic information technology planning is a tool aimed at aligning IT efforts with the organization’s business perspective by establishing, for example, policies for acquisition processes, use, and information technology management. [13].

Another view of strategic information technology planning is that it can be understood as the set of decisions and technological initiatives that support the mission, vision, and strategies of the organization over a defined period. The main purpose of information technology is to support the business itself; therefore, it must remain aligned with and focused on advancing organizational goals.

In summary, strategic information technology planning is a fundamental process that enables organizations to align their technological resources with business objectives. This process includes evaluating current and future needs, identifying technological opportunities and risks, and formulating implementation strategies.

The effective integration of strategic IT planning encourages data-based decision-making, thereby strengthening an organization’s capacity for innovation [15]. As a primary benefit, by addressing both current challenges and future opportunities, this planning enables organizations to adapt to a constantly changing organizational environment [16].

In a study conducted by Alidousti, it is stated that strategic information technology planning is effective when based on nine key factors: vision and objectives, alignment, interoperability, infrastructure, participation, equity, privacy and security, management, and organizational culture. In the present study, particular attention is given to organizational culture and management, specifically within the area of human resources [17].

2.5 Information Systems

To understand the current situation of the information technology department, it is necessary to have a clear understanding of the information system within the organization and the value it provides.

Another concept states that information systems are defined as “a set of interrelated components that collect (or retrieve), process, store, and distribute information to support decision-making and control processes within an organization” [18, p. 16].

Information systems are essential components for organizations, supporting the management and processing of data for decision-making and process optimization in an efficient manner.

These systems are composed of hardware, software, data, processes, and human resources that work together to collect, store, process, and distribute information [19]. The interaction of these components within an information system occurs in pursuit of improved productivity, cost reduction, and the provision of a competitive advantage through a clearer understanding of the market.

Each information system developed and incorporated into a business must be conceived during the strategic planning process in order to function as an active element, considering that if it is created in an earlier or later phase, it contributes only passively to the achievement of organizational objectives [14].

Finally, it can be concluded that the purpose of information systems is to improve business processes, where the use of information systems together with information technologies is intended to automate operations, streamline processes, reduce costs, shorten timeframes, promote teamwork, and improve innovation [20].

Leaving aside the strategic and technological aspects, the following section presents information regarding the National University of Costa Rica and its Huetar Norte y Caribe Regional Section in order to understand their functioning and operations.

2.6 National University of Costa Rica

The National University of Costa Rica is one of the five public universities in the country, ranking second nationwide in international rankings at position 1960, surpassed only by the University of Costa Rica, which holds position 1735 according to Webometrics [21].

The university is founded on four fundamental pillars: teaching, research, outreach, and production. It was established with the intention of providing the country with additional institutions to support education and the development of more professionals. Its mission follows the vision expressed by its first rector, Benjamín Núñez, who affirmed:

The National University of Costa Rica has expressed its mission in the following way: 'In this task, we have been concerned not so much with merely conceiving and building just one university, but giving the country a necessary university that, by effectively committing to its national reality, can serve to fulfill a historical destiny with prosperity, justice and liberty' [22].

The central headquarters is located in the province of Heredia and is known as the Omar Dengo Campus, one of the eight campuses where the university offers higher education programs. The university has almost 20,000 active students distributed across the different entities responsible for carrying out its substantive functions.

Finally, to understand its technological capabilities, the university has the Center for Technological Management (CGT), which is in charge of the university's platforms, infrastructure, and technological services. The Center for Information Management (CGI) is responsible for both the administration of technology and the development processes of the computer systems necessary for the university's substantive activities.

2.7 Regional Section Huetar Norte and Caribe

In its early phase, the Sarapiquí Campus was created under the structure of an interdisciplinary academic program named the Interdisciplinary Academic Program of the Huetar Norte and Caribbean Region, Sarapiquí Campus, within the framework of the celebration of the XXXV anniversary of the university's founding and under the principle of working toward less favored sectors [23].

According to SRHNC, "The necessary university does not center its foundations on the most developed sectors, but rather bases its work on the development of academic activities in areas whose social, economic, cultural, and environmental contributions have not been able to achieve a greater impact due to geographical limitations, as is the case of the Huetar Norte and Caribbean region" [23].

Currently, the Sarapiquí Campus offers 6 bachelor's degree programs and 3 licentiate degree programs, and for the year 2025 it has a total of 663 students, 27 administrative staff members, and 62 academic staff members.

III. METHODOLOGY

This section explains the process through which the research progressed from its initial stage to the obtaining of the final results, using the approach presented by the authors and the strategic planning proposal for information technology [5].

3.1 Type of Research

The study was conducted under a descriptive research design of the current situation of the IT department. From the perspective of data collection and analysis of results, a qualitative approach was employed.

3.2 Study Population

In this research, the study population consisted of university stakeholders involved in the services provided by the information technology department. Specifically, this included the entire student community, academic staff at the campus, administrative staff primarily linked to the leadership and management of the Regional Section, and a key actor: the head of the information technology department.

3.3 Data Collection Techniques

Given the wide range of techniques available for qualitative studies, emphasis was placed on instruments developed by the researchers themselves, allowing for the collection of information from multiple sources and enabling a deeper understanding of the phenomenon [24].

For this study, the situational phase was examined through interviews, questionnaires, and observation of the target population, with the selection of instruments based on the type of population involved (interviews for department heads and personnel, questionnaires for administrative staff, academic staff, and students, and observation of processes).

IV. RESULTS

This section compiles the information gathered during the research process, presenting the results of the application of instruments, documentary review, observation process, and analysis of each of the particularities identified in the situational assessment of the information technology (IT) department.

4.1 Current Situation of the IT Department at Sarapiquí Campus

Within the previously described context of the operations of the National University of Costa Rica and the specific case of the Sarapiquí Campus, the analysis and evaluation of the information technology department of the Regional Section are presented, specifying its functionality and areas for improvement.

Operational Functioning of the Department

The information technology department at the Sarapiquí Campus is composed of one person in a position classified as Technical Analyst in Technological Development with an assistant role, who is responsible for the following functions:

- Participating in the development and creation of networks.
- Coordinating and operating a local area computer network system.
- Assisting in the installation of equipment and wiring for institutional data network access.
- Participating in the planning and coordination of work programs.
- Conducting tests and changes in operating system versions.
- Troubleshooting hardware, software, and general network issues.
- Assisting in laboratory logistics (maintenance and quotations).
- Creating databases.
- Developing and updating website design.
- Preparing reports on activities carried out.

The working hours are regular, from Monday to Friday, 8:00 a.m. to 5:00 p.m., involving daily tasks (institutional network review, management, and responses through social media) as well as occasional duties specific to the role that arise unexpectedly.

The following figure illustrates the level of awareness among the administrative staff, teaching staff, and student population at the Sarapiquí Campus regarding the services offered by the IT department.

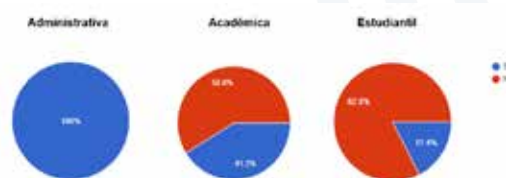


Fig. 2. Knowledge of IT Services by Administrative, Academic, and Student Population.

As shown in Fig. 2, *Knowledge of IT services by Administrative, Academic, and Student Population*, the majority of academic staff and students are unaware of the functions and services offered by the department, while only the administrative staff demonstrate familiarity with the tasks performed.

4.1.2 Processes

Within the operational functioning of the IT department, although there is no formal written definition of processes, the following can be identified.

Table 1. IT Processes.

Process	Description
Acquisition of Technological Equipment	All technology equipment purchases are subject to the expert judgment and recommendations of the IT department.
Definition of the Campus' own information architecture	In its operations, and outside the general administrative systems of the National University of Costa Rica, the Campus continues to rely primarily on paper-based communication, where there are no established schemes, repositories, or specific locations for the storage of important information, but rather a dependence on the organizational practices of each individual member.
Management of IT projects	There is no formal guide for the execution of IT projects; instead, they are carried out according to urgency, with limited documentation and follow-up.
IT risk evaluation and management	Knowledge of the risks associated with the functions and their possible solutions rests with the IT manager, since there is no documentation of the risks specific to the department's management. As there is no risk documentation, no evaluations are conducted to verify the reduction of their impact.
University community relations management	Maintain clear communication channels and mechanisms with the university community as a whole.

4.1.3 Infrastructure

According to the organizational structure of the Sarapiquí Campus, the IT department must report its work to the head of administrative services of the Regional Section, making it one of the few technical areas of the university that reports to personnel outside its field.

In the context of the Regional Section itself, the information technology infrastructure can be understood as comprising the desktop environment (computers, printers, photocopiers, etc.), the applications environment (software used and licenses), and the network and telecommunications environment (interconnectivity equipment, licenses, and bandwidth).

4.1.3.1 Desktop Environment

These are the components that form part of the technological equipment and end-user device environment, which, in the case of the Sarapiquí Campus, include the following:

Table 2. Desktop Environment.

Devices	Amount	Description
Computers	125	DELL and HP computers, which are part of the Campus's two laboratories, computers assigned to the administrative staff for their respective duties, as well as a "mobile laboratory" consisting of computers used in classrooms by students or loaned for the completion of academic tasks.
Printers	5	Printing services for the administrative staff, academic staff, and student population.
Photocopiers	2	Photocopying services for the administrative and academic staff; there is no specific equipment available for student use.
Interactive whiteboards	11	Installed in each classroom to support teaching processes.
Projectors	19	Installed in each classroom to support teaching processes. Other projectors are available for loan for activities such as CONARE's induction tours to schools.
Speakers	3	Loan equipment.
IP Phones	13	Devices in administrative offices and academic rooms to enable internal and external communication.
Robotics kits	16	Kits for teaching robotics.
External hard disk	1	Information backup.

For the maintenance or replacement of technological equipment, a budget request is submitted with the corresponding justification on the dates established by the university. Currently, preventive and corrective inspections of the equipment are carried out (at least once every six months), and the inventory is reviewed once a year.

It is important to note that the total equipment data available at the time of the interview was not accurate, since the inventory process was still ongoing, and that the data presented in Table 2, *Desktop Environment*, was compiled using information provided by the IT manager and corroborated through observation.

4.1.3.2 Application Environment (Information Systems)

The National University of Costa Rica has two different operating system solutions for its devices: a free software option with its own customized distribution called UNAX, which sought to facilitate the transition from proprietary software to free software. However, the migration, at least at the Sarapiquí Campus, was not successful, and gradually the devices returned to using proprietary operating systems.

With respect to proprietary operating systems, Windows 11 is installed in the laboratories and on the equipment of the Campus.

As evidenced by the instruments applied to the student, administrative, and academic populations, the issue of applications is not a crucial factor in their assessment of the IT department, and both the operating systems and administrative systems provide adequate support for the work carried out by these groups.

4.1.3.3 Server Environment

The server environment is provided by the Center for Technological Management and not by the IT department of the Sarapiquí Campus; therefore, the storage of its websites, applications, and documentation is housed at the Central Campus of the National University of Costa Rica. Physically, there is a virtual server hosted at DTIC for the exclusive use of the applications and requirements of the Sarapiquí Campus.

4.1.3.4 Network and Telecommunications Environment

This is understood as the set of components that form part of the network and telecommunications environment, which are presented in Chart 3.

Table 3. Network and Telecommunications Environment.

Devices	Amount	Description
Router	1	Routing necessary for internet outlets, configured and managed by the Technology Management Center; the IT manager is responsible for conducting inspections and reporting incidents.
Switch	7	Supports internal routing, located in different racks to cover the geographical area of the Campus.
Standing rack	1	Located in the IT department office.
Aerial rack	3	Frame that protects the network intermediary equipment, located in different locations
Firewall	1	Limits the entry and exit of traffic of doubtful origin.
Access point	9	Installed in the main esplanade, auditorium, guard-house, and residences to provide wireless connectivity.

As shown in Chart 3, *Network and telecommunications environment*, the university has adequate equipment for the development of its functions in accordance with the number of users and geographical distribution. In addition, it has a fiber optic link between the Campus and the Central Campus with 1 gigabit internet capacity.

It is pertinent to point out, based on the observation of the physical facilities of the IT department, that using the office space of the department's personnel as an equipment inventory warehouse and telematics room for intermediary network devices does not constitute good practice.

4.1.4 Projects

In addition to the existing processes within the department, there are a series of projects necessary for the optimal fulfillment of its functions and for addressing stakeholders' emerging needs, with the main projects listed in Chart 4.

Table 4. IT Projects Perceived by Stakeholders.

Project	Description
Clarification of services	Each service offered by the department is provided at the request of one of the stakeholders; however, there is no documentation of the services offered, nor any alignment with a corresponding reference framework.
Project management methodology	Each project is carried out according to its need and urgency; however, there is no documentation process for its evaluation, and therefore there are no business cases or follow-up procedures for the projects being executed.
Process documentation	Each process is developed according to previous experience; however, the processes are not documented in a way that would allow another staff member to perform them.
Modernization of technological equipment	A large portion of the equipment on the Campus has already reached the end of its useful life cycle, and the IT department is required to carry out the acquisition process for its replacement.

The projects presented in Chart 5 reflect the needs identified for the next two years according to the perceptions of the stakeholders. To implement these initiatives, close coordination with the central administration is required in order to support the proposals presented in the following section.

4.1.5 Communication with Headquarters

The immediate superior authority of the technology department is the head of administrative services, with whom, according to the information gathered through interviews with both actors, there is stable and trusting communication. From the perspective of the department head, the work carried out by the department is validated.

However, with respect to Campus management, communication is weaker and less timely regarding strictly work-related matters, due to requirements in tasks that, from the perspective of the Regional Section management, have not been fully or optimally developed.

In addition, the vision of the Regional Section management toward the IT department reflects a relative absence of expectations regarding collaboration in innovation projects, since the established role does not contemplate such extraordinary functions.

4.1.6 Services

The IT department, through its processes and daily operations, carries out a series of activities and services that support the substantive work of the Sarapiquí Campus, as shown in the following chart.

Table 5. IT Services.

Service	Description
Telematics equipment support	Daily diagnostics of the equipment, in which connectivity between devices and external networks is verified.
Support to the Campus connectivity network	In case of minor incidents or local failures, a first review is made from the department, validating that they are not user or equipment errors.
Support for Campus computer equipment	Overhaul and repair of technological equipment on campus, such as computers, monitors, printers, photocopiers, network equipment, student equipment, among others.
Laboratory maintenance	Review conducted once per semester of the equipment available in the Campus laboratories, including its hardware and software operation, equipment replacement, among other aspects.
Acquisition of technological equipment	In charge of approving and advising on purchases related to technological equipment.
Equipment storage and loan services	To maintain equipment inventoried and stored for safekeeping and timely loan according to the needs of the university population.
Training on technological tools and platforms	As a consequence of the COVID-19 pandemic, training on technological tools for first-year students, as well as on the use of administrative, academic, and educational platforms and administrative systems for support and control purposes, should be provided.
Advice on the use of equipment and connectivity	Group training and/or guidance on the use of sound equipment, microphones, interactive whiteboards, and connectivity for users invited to the Campus.
Support with audio-visual and multimedia equipment	The IT department is responsible for providing sound support for activities held on the Campus; however, other personnel should be trained to perform these tasks, since the department is left unattended during such activities. Likewise, support materials (documents and videos) should be developed for the use of the interactive whiteboards.
Management of digital communication channels	Greater flexibility and a larger number of channels should be established as mechanisms for the dissemination and presentation of the activities carried out to the university community in general.

Derived from chart 6 IT Services, there are many services that are broken down into subtasks, each requiring effort and time, but they are not being visualized by university stakeholders. Therefore, their documentation and offering are imperative.

4.1.7 SWOT Analysis

The IT department at the Sarapiquí Campus has 10 years of experience in carrying out its functions; however, to date, no audits or reviews have been conducted to support the evaluation and development of the department. Therefore, a SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis is carried out to provide a clearer understanding of the support that information technology offers to the university campus. These aspects will also be considered in the recommendation phase.

- Strengths: These consist of favorable aspects related to the department's performance, including:
- Availability and willingness of staff to support the request and development of services.
- Level of influence in decision-making regarding technological solutions and equipment.
- Experience of the department's human resources.
- Opportunities: These are identifiable elements that can enhance both the capabilities offered by the department and the perception of the stakeholders involved in the university community, thereby positively supporting its mode of operation. The identified opportunities are:
- Training administrative and academic staff for tasks that do not require extensive technical knowledge.
- Knowledge management of the services performed (historical documentation).
- Implementation of a traceability system for its services and tasks.
- Weaknesses: These represent critical points in management that require improvement, as they pose inconveniences and vulnerabilities for the department. The following can be identified:
- Student perception of the services offered by the department.
- The perception of the department's "absence" by the Campus management.
- Lack of tools for storing the inventory of equipment managed by the department.

Threats: Finally, threats represent weaknesses or points of concern that should be considered potential elements of change and that may affect the department's operations. The main threat identified relates to the administration's commitment to supporting the projects developed by the IT department, since without such support they cannot be implemented.

Simultaneously, another potential threat is a change in leadership during the strategic IT planning process, since if the ongoing project does not have a solid institutional foundation or is not considered relevant by the new administration, it may be neglected.

Lastly, an additional threat is the possible outsourcing of the department's services through the subcontracting of personnel or another entity to carry out its functions. However, due to the collective agreement, internal university policies, and the university union, this option is not considered highly viable.

4.1.8 Budget

The Sarapiquí Campus has a budget allocated by the National University of Costa Rica, supervised by the Deputy Rector within the university's organizational structure. There are no specific budget allocations assigned to each department; therefore, according to information gathered through interviews with the IT director and Campus management, the IT department does not have a dedicated budget for investments or operational expenses.

Any project initiative must be submitted first to the head of services and then to the Campus management for evaluation and possible inclusion in the Regional Section's budget plan for the following year. Alternatively, it may be supported through the reallocation of unspent funds from other areas, although this is a less likely option.

Thus, the IT budget is tied to the requests and innovation initiatives of the department, relying on prior anticipation of needs rather than on an exclusive fund for expenses. Therefore, the department must have a clear understanding of the projects to be executed, develop the corresponding business cases, and determine the budget required for their implementation, all of which must be presented to the immediate management in order to assess their feasibility.

For this strategic plan, the proposed projects require a low budget investment for two reasons. First, at the beginning of the second semester of 2024, there are no resources allocated to the department (requests should have been submitted before September 2023). Therefore, the nature of the projects is focused on training, standardization, and management, resulting in a low economic impact that can be covered with the Campus's own resources.

Secondly, managing smaller projects with high success rates allows the department to build trust and improve its perceived image among stakeholders.

4.1.9 Gap Analysis

After applying the instruments to the involved stakeholders, an analysis of the current situation and the path toward the desired situation was conducted, revealing misalignments in aspects such as:

- There is a negative perception of the quality of services offered by IT, mainly expressed by the student population.
- No follow-up on service requests or IT projects, as there is no documentation or traceability of the actions requested by the involved populations.
- Standardization and documentation, both physical and digital, of equipment request forms and service tickets, as there is no formal mechanism.
- No clear identification or documentation of the department's processes.
- Knowledge management for the development of functions, as there is a risk associated with the concentration of information in a single person—a critical resource.
- Use of software programs for the development of daily administrative tasks (e.g., a service desk).
- Skills in programming or researching tools to develop solutions for specific campus needs (e.g., programming a reservation application for the use of official vehicles).
- Proactivity in the IT department's functions and its role as a strategic unit within the organization, moving beyond a solely operational function.
- Furthermore, after conducting the gap analysis for the IT department of the National University of Costa Rica, Sarapiquí Campus, several points of interest became evident:

- Although most of the required services by the department exist on the campus, essential stakeholders are unaware of these facilities.
- Services are developed according to the availability of the department's human resources, but requesters have no way to verify the status of their requests or receive timely responses.
- There is no project management that enables their completion and traceability.
- The number of tasks performed by the department, including routine job activities and service activities, results in a workload that exceeds the possible working hours.
- The availability and openness of the department's human resources are adequate for their performance.
- There is a considerable number of activities covered by the IT department that could be handled by other campus personnel, reducing the workload.
- The perception of the department's work varies among stakeholders, with administrative staff giving the highest ratings and students the lowest (with a generally low average rating).

V. CONCLUSIONS

After carrying out the research process and analyzing the findings to finalize the proposal regarding the current situation of the IT department at the National University of Costa Rica, Sarapiquí Campus, a series of key points for reflection can be highlighted:

Involve the IT department staff as key stakeholders in the development of the institutional strategic plan.

The department has trained personnel to carry out its functions, and although the workload may seem excessive in some cases, benchmarking shows a similar number of resources per educational entity. This is due to service demands peaking during academic periods and decreasing considerably once these periods end.

It is imperative to establish a constant communication channel and regular reporting from the IT department to the Campus management, demonstrating and providing evidence of the efforts carried out. In this way, the management can become a strategic ally and sponsor of the department's projects.

The image of the IT department regarding the services it offers to university stakeholders is not encouraging. As a strategy for improvement, it is proposed to carry out the projects established in the proposal, gradually making the department's work more visible, applying user satisfaction instruments, and validating this behavior.

It is recommended to carry out the necessary procedures during the stipulated periods for requesting budget funds, with the corresponding justification and business cases for each project, in order to achieve a greater impact and support solutions that require more than effort and goodwill alone.

Acknowledgment

In general, we would like to thank the National University of Costa Rica (UNA), and in particular the staff of the IT department, teaching staff, administrative staff, and students of the Sarapiquí Campus for their collaboration and contributions to the research. We also extend our gratitude to the academic colleague from Campus Coto, Angie Gómez Mora, for her collaborative work in the study.

Competing interests

The authors declare no competing interests.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Steven Cruz-Sancho <https://orcid.org/0000-0001-5549-4990>

Angie Gómez-Mora <https://orcid.org/0009-0005-2286-3661>

Olivier Blanco-Sandí <https://orcid.org/0009-0007-1182-4856>

Rebeca Elizondo-Pizarro <https://orcid.org/0009-0006-0097-9273>

REFERENCES

- [1] R. L. Daft, *Essentials of Organization Theory and Design*. Nashville, TN, USA: South-Western Educational Publishing, 2001.
- [2] J. Q. Castellanos and M. B. Arguello, "Liderazgo y cultura. Influencia en la eficiencia organizacional," *Revista Publicando*, vol. 5, no. 14(2), pp. 286–302, 2018. [Online]. Available: <https://core.ac.uk/reader/236644410>
- [3] P. Sparrow and C. Cooper, "Organizational effectiveness, people and performance: new challenges, new research agendas," *Journal of Organizational Effectiveness: People and Performance*, vol. 1, no. 1, pp. 2–13, 2014.
- [4] R. M. Shiferaw and Z. A. Birbirsa, "Digital technology and human resource practices: A systematic literature review," *Heliyon*, vol. 11, no. 2, e41946, 2025. doi:10.1016/j.heliyon.2025.e41946.
- [5] A. Cassidy, *A Practical Guide to Information System Strategic Planning*. New York, NY, USA: Taylor & Francis Group, 2006.
- [6] J. C. Sandí Delgado and M. A. Cruz Alvarado, "Propuesta metodológica de enseñanza y aprendizaje para innovar la educación superior," *InterSedes*, vol. 17, no. 36, pp. 153–189, 2016. doi:10.15517/isucr.v17i36.27100
- [7] C. A. Hernández, M. G. Gómez, and M. Balderas, "Inclusión de las tecnologías para facilitar los procesos de enseñanza-aprendizaje en ciencias naturales," *Revista Electrónica Actualidades Investigativas en Educación*, vol. 14, no. 3, pp. 1–19, 2014. [Online]. Available: <https://www.redalyc.org/articulo.oa?id=44732048010>
- [8] R. Saavedra, L. Castro, O. Restrepo, and A. Rojas, *Planificación del Desarrollo*. Bogotá, Colombia: Editorial Fundación Universidad de Bogotá, 2014. [Online]. Available: shorturl.at/lpX29

- [9] S. Robbins and M. Coulter, *Administración*. Ciudad de México, México: Pearson Education, 2005.
- [10] M. E. Porter, "What Is Strategy?," *Harvard Business Review*, 1996.
- [11] H. Koontz, H. Weihrich, and M. Cannice, *Administración: Una perspectiva global*. Ciudad de México, México: McGraw-Hill, 2012.
- [12] L. Serrano Cárdenas, E. Bravo Ibarra, and B. Amante García, "Una metodología innovadora para la planeación estratégica en las universidades," *INGE CUC*, vol. 9, no. 1, pp. 43–63, 2013.
- [13] J. Clempner and A. Gutiérrez, "Administración y ejecución de un plan estratégico de tecnología de información," *Revista Digital Universitaria*, vol. 3, no. 1, pp. 1–16, 2002. [Online]. Available: <http://www.revista.unam.mx/vol.3/num1/art1/index.html>
- [14] R. J. Jacinto Jáuregui and J. P. Santos Fernández, "Planeamiento estratégico de tecnologías de la información y su impacto en la mejora de la gestión de servicios de TI en la Facultad de Ingeniería de la UNT," Bachelor's thesis, Univ. Privada del Norte, Trujillo, Perú, 2018. [Online]. Available: shorturl.at/jDY13
- [15] J. Peppard and J. Ward, *The Strategic Management of Information Systems: Building a Digital Strategy*. Hoboken, NJ, USA: Wiley, 2016.
- [16] J. Luftman, K. Lyytinen, and T. B. Zvi, *Strategic Alignment of Information Systems: Building Capabilities for Sustained Success*. New York, NY, USA: Routledge, 2022.
- [17] S. Alidousti and F. Sahli, "National and sectoral information technology planning: a systematic literature review," *International Journal of Public Sector Management*, vol. 37, no. 4, pp. 465–485, 2024. doi:10.1108/IJPSM-09-2023-0286.
- [18] K. C. Laudon and J. P. Laudon, *Sistemas de Información Gerencial*, 14th ed. Ciudad de México, México: Pearson Education, 2016.
- [19] P. Beynon-Davies, *Information Systems: An Essential Guide to Business and Technology*. London, U.K.: Palgrave Macmillan, 2023.
- [20] E. E. V. Encalada, R. A. R. Lozano, F. G. Oscco, and F. D. M. S. Aguirre, "Sistemas de información como herramienta para reorganizar procesos de manufactura," *Revista Venezolana de Gerencia*, vol. 24, no. 85, 2019. [Online]. Available: <https://www.redalyc.org/articulo.oa?id=29058864015>
- [21] C. S. Científicas, "Ranking web of universities," *Webometrics.info*, Feb. 10, 2023. [Online]. Available: https://www.webometrics.info/es/Latin_America_es/Costa%20Rica
- [22] Universidad Nacional de Costa Rica (UNA), "Universidad Nacional de Costa Rica," *Portal de Transparencia UNA*, 2019. [Online]. Available: https://www.transparencia.una.ac.cr/index.php?option=com_content&view=article&id=297&Itemid=741
- [23] SRHNC, "Historia de la Sección Regional Huetar Norte y Caribe," *Sitio web de la SRHNC*, Heredia, Costa Rica, 2022. [Online]. Available: <https://www.srhnc.una.ac.cr/>
- [24] R. Hernández-Sampieri and C. Mendoza Torres, *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. Ciudad de México, México: McGraw-Hill, 2018. ISBN: 978-1-4562-6096-5.



AVANCES EN EL RECONOCIMIENTO DE PLACAS VEHICULARES CON VISIÓN COMPUTACIONAL Y MACHINE LEARNING EN COSTA RICA

T. Bermúdez

*Universidad Cenfotec, Francisco Peralta, San José
tbermudezm@ucenfotec.ac.cr*

J. Hidalgo

*(Universidad Cenfotec, Francisco Peralta, San José
jhidalgou@ucenfotec.ac.cr)*

L. Naranjo

*(Universidad Cenfotec, Francisco Peralta, San José
lnaranjo@ucenfotec.ac.cr)*

RESUMEN

El presente artículo documenta el desarrollo de un sistema de visión computacional y aprendizaje automático para el reconocimiento de placas vehiculares en Costa Rica, un área poco investigada localmente. Utiliza el algoritmo YOLOv5 y técnicas de OCR para detectar vehículos y procesar texto en tres etapas: detección vehicular, detección de placas y reconocimiento de caracteres. Se entrenaron dos modelos, con lo cual se logró un 77 % de precisión en detección de vehículos y 96 % en detección de placas. Además, se creó un dataset experimental con 729 placas costarricenses, donde el modelo alcanzó un 69 % de verdaderos positivos. Esta solución representa un avance significativo para aplicaciones de seguridad vial y propone una base sólida para futuros desarrollos. Se recomienda ampliar el dataset nacional para mejorar el rendimiento.

Palabras clave: Aprendizaje automático, conjunto de datos vehiculares, placas vehiculares, reconocimiento de caracteres, reconocimiento de matrículas, seguridad vial, visión computacional.

ABSTRACT

This paper documents the development of a computer vision and machine learning system for license plate recognition in Costa Rica, a field with limited local research. It employs the YOLOv5 algorithm and OCR techniques to detect vehicles and process text in three stages: vehicle detection, license plate detection, and character recognition. Two models were trained, achieving 77% accuracy in vehicle detection and 96% in plate detection. Additionally, an experimental dataset of 729 Costa Rican license plates was created, where the model achieved a 69% true positive rate. This solution represents a significant advancement for road safety applications and provides a solid foundation for future developments. Expanding the national dataset is recommended to improve performance.

Keywords: Character recognition, computer vision, license plate recognition, license plates, machine learning, road safety, vehicle datasets.

I. INTRODUCCIÓN

En la actualidad, la seguridad vial se ha convertido en una preocupación primordial en muchas sociedades, y Costa Rica no es una excepción. La necesidad de implementar soluciones innovadoras que contribuyan a prevenir siniestros y mejorar la identificación de vehículos implicados en eventos delictivos ha llevado al desarrollo de proyectos basados en tecnologías informáticas avanzadas, como la visión computacional y el machine learning.

Bajo este sentido, la investigación tiene como objetivo principal desarrollar un sistema de reconocimiento de placas vehiculares en vías públicas de Costa Rica. De modo que se enmarca en un enfoque alternativo, orientado hacia la demostración, pruebas y análisis práctico, se prioriza el desarrollo funcional del sistema, y el cumplimiento de los objetivos de investigación, estableciendo un avance en esta área.

La relevancia de este trabajo radica en la ausencia de implementación de sistemas de detección de vehículos en vías públicas en Costa Rica, situación que ha generado la necesidad de investigar y desarrollar un proyecto que aborde esta problemática de manera integral. Esto cobra especial importancia ante estadísticas preocupantes de muertes en sitio por accidentes de tránsito, ya que, para el año 2024, se registraron 513 personas fallecidas en el lugar del accidente, lo que representa una tasa de mortalidad de 9,93 por cada 100 000 habitantes, esto deja en evidencia la urgencia de fortalecer la seguridad vial en el país [1]. Asimismo, según el Organismo de Investigación Judicial [2], en 2020 se contabilizaron 3165 denuncias por robo de vehículos, mientras que en 2023 esta cifra aumentó a 4562, ello refleja un incremento significativo en este delito [2]. Esta alta incidencia y el aumento sostenido de casos, junto con los riesgos persistentes en carretera, resaltan la imperiosa necesidad de implementar sistemas de detección que contribuyan a la reducción de estos delitos y al mejoramiento de la seguridad en las vías públicas de forma preventiva.

El enfoque alternativo, basado en el paradigma pragmático [3], responde directamente al objetivo de desarrollar un sistema funcional y práctico para Costa Rica, donde no existen implementaciones previas. Al priorizar el desarrollo funcional y un producto de software operativo, se asegura la creación de un sistema tangible y utilizable. Asimismo, la orientación hacia la demostración, las pruebas y el análisis práctico garantiza que el sistema sea evaluado rigurosamente en su capacidad para reconocer placas vehiculares en el contexto costarricense, con lo cual cumple con los objetivos de investigación y aborda la necesidad de una herramienta innovadora para fortalecer la seguridad vial en el país. Este enfoque en particular se basa en tres dimensiones fundamentales:

- Epistemológica: Centrada en la posición de los investigadores como observadores en la búsqueda de comparar diversas técnicas y arquitecturas de reconocimiento de matrículas.
- Ontológica: Enfocada en la conceptualización de los elementos clave para entender la estructura del sistema ALPR (Automatic License Plate Recognition).
- Axiológica: Estudia los valores y define criterios de eficiencia y precisión para la evaluación de los resultados del sistema.

La investigación se apoya en el uso de técnicas de visión por computador y aprendizaje automático, basándose en el análisis de conceptos relevantes para el reconocimiento de placas vehiculares. El

sistema resultante ha sido sometido a una evaluación rigurosa, considerando factores como la velocidad de procesamiento y la precisión en la identificación de placas.

En el apartado 2, se expone una revisión de literatura en la que se abordarán las bases teóricas sobre sistemas ALPR, dichas bases guiarán los métodos a seguir en el capítulo 3 para la elaboración del sistema. En el apartado 4, se abordarán y expondrán los resultados experimentales encontrados en el desarrollo. Finalmente, en el apartado 5 se entrará en discusión sobre el avance alcanzado junto con las limitaciones técnicas del sistema para la detección de placas vehiculares en vías públicas, además, se presentarán puertas de investigación en campos similares y trabajos a futuro.

II. REVISIÓN DE LITERATURA

En esta sección los autores detallan una revisión de literatura, enfocada en formar criterio técnico sobre el machine learning y visión computacional para el proceso de detección de placas vehiculares. Siguiendo la metodología planteada por Biolchini et al. [4], se realizó la selección de los trabajos candidatos y su posterior revisión, bajo la cual se eligieron artículos y artículos académicos para formar parte de las bases teóricas y técnicas de la investigación. Los artículos seleccionados contienen una gran cantidad de conceptos, información, modelos, técnicas y recomendaciones para la construcción de un sistema LPR eficiente y robusto.

Como punto de partida, Kulkarni et al. [5] abordaron una revisión de los sistemas de reconocimiento automático de matrículas vehiculares (ALPR), en la cual describieron tres procesos principales: extracción de matrícula, segmentación de caracteres de matrícula y reconocimiento de la matrícula, adicionalmente a estas tres etapas referentes a la arquitectura del sistema ALPR, Du et al. [6] definieron una cuarta etapa anterior a estas, denominada adquisición de imágenes, prestando especial atención a los datasets empleados en el aprendizaje automático.

Para cubrir el tema de adquisición de imágenes, Björklund et al. [7] emplearon un dataset constituido por imágenes sintéticas para el entrenamiento de un modelo de detección, el cual evaluaron posteriormente en un conjunto de datos de imágenes de placas vehiculares del mundo real, donde se evidencia efectividad para reconocer con precisión placas de matrícula en diversos escenarios. Asimismo, Wang et al. [8] utilizaron también un

dataset sintético y compararon la precisión y la eficacia en la detección y reconocimiento con matrículas vehiculares reales y se obtuvieron buenos resultados en estos dos aspectos.

En contraposición a esta técnica, Laroca et al. [9] entrenaron con un dataset público para ALPR con más de 4500 imágenes de matrículas vehiculares, se alcanzaron mejores resultados comparados con sistemas comerciales y datasets existentes. Similar a estos autores, Amirgaliyev et al. [10] implementaron y probaron un método ALPR, en un dataset de 1470 imágenes, donde resultó que había una mayoría de falsos positivos antes de la verificación.

Para la siguiente etapa, referente a la extracción de placas vehiculares, Xu et al. [11] presentaron un método de detección de placas vehiculares, utilizando OpenCV, comparando y evaluando el rendimiento de varios algoritmos de detección de matrículas. Entre ellos evaluaron el método de detección de bordes de Canny, el método de detección de características de Haar, el método de detección de características de LBP y el método de detección de características de HOG, destacando este último como el método con mejor precisión y velocidad de detección. Por otro lado, Yu et al. [12] propusieron un método para ubicar placas vehiculares mediante la transformación de Wavelet y el análisis de descomposición modal empírica (EMD). Los resultados demuestran su efectividad para ubicar con precisión las placas de matrícula en varios escenarios.

Lin et al. [13] propusieron un método para la ubicación de placas vehiculares combinando la detección de bordes y la morfología matemática. El algoritmo fue diseñado para reducir la interferencia del ruido en las imágenes y aumentar la precisión en tiempo real. De forma similar, Giannoukos et al. [14] presentaron un algoritmo llamado Operator Context Scanning (OCS), que mejora la velocidad de procesamiento de videos e imágenes de alta definición en el reconocimiento de matrículas de vehículos en tiempo real en más del 250 % sin afectar la precisión.

Ahn et al. [15] describieron el diseño de un algoritmo de detección de zonas de interés (ROI) de placas vehiculares para mejorar su reconocimiento, empleando una combinación de técnicas de detección de bordes y análisis de características para identificar las ROI propensas a contener una placa de matrícula.

La capacidad de detección de ROI es un paso clave al que prestarle especial atención, esto se ve reforzado por el artículo de Bakhtan et al. [16] en el que, examinando los pasos claves para la extracción de la región de interés, concluyen que los métodos basados en redes neuronales y aprendizaje profundo se destacan por su precisión y eficiencia.

Sadique y Haque [17] revisaron varios trabajos sobre técnicas de detección y reconocimiento de placas vehiculares, incluyendo métodos basados en características, aprendizaje profundo y redes neuronales, de igual forma, se concluye que estos son más precisos y eficientes, que métodos previos. Pustokhina et al. [18], por ejemplo, presentaron un método ALPR utilizando una combinación de clustering K-means óptimo y redes neuronales convolucionales (CNN), teniendo como objetivo mejorar precisión y eficiencia en sistemas ALPR para el transporte inteligente. Los autores realizaron experimentos en un conjunto de datos de imágenes de matrículas y lograron altas tasas de precisión, superando otros métodos. Pu et al. [19] presentaron un método de CNN de detección de matrículas vehiculares extrayendo y fusionando características de bajo y alto nivel de las matrículas. El resultado del modelo fue eficiente en variedad de entornos complicados y logró una detección precisa y en tiempo real.

Los avances contemporáneos en cuanto al uso de deep learning para la detección de placas en imágenes proporcionan una guía para el desarrollo del sistema ALPR y la arquitectura a utilizar para su implementación en diversos entornos. Como evidenciaron Menon y Omman [20], con su sistema de reconocimiento de placas de matrícula, empleando técnicas de procesamiento de imágenes y aprendizaje automático, los autores probaron placas vehiculares de India y España, demostrando una alta precisión.

En este punto, también se aborda el trabajo de Shobha y Deepu [21], los cuales llevaron a cabo una revisión de métodos de detección de vehículos basados en visión, incluyendo enfoques de aprendizaje profundo como CNN, RCNN y DNN, además de analizar los desafíos asociados con la detección y clasificación de vehículos. Previamente, Kuo et al. [22] presentaron un sistema ALPR conformado por dos subsistemas para detección de automóviles y extracción de placas. Para el primero, usaron reglas de decisión de cantidad mínima de movimiento para encontrar la ubicación de los automóviles en la pantalla, para el segundo utilizaron características de las placas y un algoritmo de búsqueda para extraer las placas.

Palanivel et al. [23] propusieron un sistema para detectar placas vehiculares en situaciones difíciles, como en condiciones de luz distorsionada, alta o baja, y polvo, usando Faster R-CNN, segmentación de cuadros, interpolación de imágenes y reconocimiento óptico de caracteres (OCR). Este enfoque para la detección de caracteres destaca para la etapa de extracción de los números de matrícula vehicular. Este paso previo de detección vehicular se presenta como una etapa extra para realizar una depuración de datos previo a la extracción de las zonas de interés, los cuales pueden brindar una ventaja a la hora de descartar falsos positivos.

Por tanto, basados en este conjunto de artículos, se tomaron las bases para definir los métodos empleados para la creación de un sistema ALPR. El planteamiento es posible gracias a las múltiples intervenciones de autores internacionales en este ámbito, guiándonos con base en su experiencia y recomendaciones. Estos métodos, los cuales se detallan a continuación, parten del diseño metodológico, la recolección de datos y los ambientes y tecnologías de desarrollo.

III. MÉTODOS

En este apartado, se detallan los elementos críticos bajo los cuales se elabora la construcción del sistema.

Tipo de investigación

El tipo de investigación abordado para la selección de métodos o técnicas es evaluativo, derivado de los hallazgos en la revisión de literatura previamente planteada.

Alcance investigativo

Dado el tipo de investigación que se está realizando, esta se circunscribe a los siguientes alcances investigativos:

- **Exploratorio**
La investigación exploratoria busca familiarizar al investigador con un tema no abordado antes, novedoso o poco estudiado [24]. En nuestro caso, si bien se han efectuado investigaciones sobre reconocimiento de matrículas vehiculares alrededor del mundo, no se encontraron estudios o aplicaciones para Costa Rica.
- **Descriptivo**
Las investigaciones descriptivas se definen como “útiles para analizar cómo es y cómo se manifiesta un fenómeno y sus componentes” [24]. Este enfoque resulta útil para el análisis e interpretación de experimentos que se realicen a lo largo de la investigación.

Enfoque

Como parte del apartado de marco metodológico, se plantea el uso del enfoque alternativo para el desarrollo del reporte de investigación. Lo anterior dada la necesidad de generar una investigación orientada en la demostración, pruebas y análisis. Como resultado del uso de este enfoque, se presentarán tres dimensiones de investigación fundamentales abordadas de forma detallada en este reporte:

Dimensión epistemológica

Esta dimensión abarca la posición de los investigadores frente al fenómeno o acontecimiento de estudio. En este caso, dada la naturaleza de una investigación evaluativa de las principales técnicas y arquitecturas usadas en las fases de un sistema ALPR, los investigadores van a asumir un rol de observadores, en la búsqueda de comparar diversas hibridaciones de técnicas, sus resultados y precisión.

Dimensión ontológica

Desde el punto de vista ontológico, se parte de la conceptualización de los conceptos relevantes para el entendimiento de una estructura LPR, derivados de la profundización por medio de la revisión de literatura. La Fig. 1 representa los conceptos más relevantes para este fin:



Fig. 1. Ontología de estructura LPR (Elaboración propia).

Dimensión axiológica

La dimensión axiológica corresponde al estudio de los valores, definiendo escalas de valores para evitar el uso de conceptos cuya definición llegue a ser muy ambigua o vaga en el contexto de la investigación. Los resultados de las estructuras planteadas deben ser sometidos a una evaluación de acuerdo con las rúbricas planteadas en la Tabla 1.

Tabla 1. Criterios para clasificar eficiencia y precisión (Elaboración propia).

Nombre	Descripción	Evaluación
Eficiencia	Velocidad del sistema o método a la hora de su ejecución en tiempo real. Se mide en segundos.	Altamente eficiente Se considera altamente eficiente un sistema o método que produzcan su resultado en un tiempo menor de 5 segundos.
		Medianamente eficiente Se considera medianamente eficiente un sistema o método que llegue a un resultado en un tiempo de 5 a 8 segundos.
		Poco eficiente Se considera poco eficiente un sistema o método que llegue a un resultado en tiempo mayor a 8 segundos.
Precisión	Porcentaje de procesamiento de datos exitoso del sistema o método. El porcentaje se calcula respecto a la cantidad de datos y a la cantidad de aciertos.	Altamente preciso Es altamente preciso el sistema o método con un porcentaje de aciertos mayor o igual al 95 %.
		Preciso Es preciso el sistema o método con un porcentaje de aciertos mayor o igual a 90 % y menor a 95 %.
		Medianamente preciso Es medianamente preciso el sistema o método cuyo porcentaje de aciertos sea mayor o igual a 70 % y menor a 90 %.
		Poco preciso Es poco preciso el sistema o método con un porcentaje de aciertos menor al 70 % y mayor a 60 %.
		Apenas preciso Es apenas preciso el sistema o método con un porcentaje de aciertos menor a 60 %.

Dataset

A pesar de que este tema de investigación ha sido abordado por parte de diferentes autores internacionales, no fue posible encontrar un conjunto de datos de placas vehiculares exclusivamente costarricenses y de acceso al público para el uso de datos a nivel nacional. Debido a esta situación,

se han utilizado datasets de origen internacional, siguiendo la etapa de la adquisición de imágenes para el aprendizaje automático recomendada en [6]. Estos se detallan a continuación como el material de entrenamiento de los modelos desarrollados.

Dataset de detección de placas

El dataset empleado para el entrenamiento del modelo destinado a la detección de placas vehiculares e identificación de ROI se denomina "Automatic Number Plate Recognition" y fue subido a la plataforma Kaggle por el usuario Aslan Ahmedov, contando con una calificación de 10 en usabilidad dada por la plataforma. Este dataset contiene 453 archivos, compuestos de 228 imágenes en formato JPEG en distintas resoluciones y 225 archivos en formato XML, donde se incluyen anotaciones de las imágenes. Las anotaciones se proporcionan en el formato PASCAL VOC, un estándar en el campo de la detección de objetos y localización, ampliamente utilizado por diferentes editores y herramientas de anotación para entrenar modelos de aprendizaje automático. Dichas anotaciones incluyen información como cuadros delimitadores (bounding box) de las matrículas de vehículos presentes en las imágenes, detalles de la imagen, clases, rotación y otros datos relevantes. Las zonas de interés se encuentran etiquetadas bajo la clase "license_plate" para diferenciar la placa vehicular del resto de la imagen.

Dataset de detección de vehículos

Para el entrenamiento del modelo de detección de vehículos utilizando YOLOv5, se ha empleado un dataset específico a partir del trabajo previo de Anhalu en su repositorio "vehicle-object-detection-using-YOLOv8", dataset originalmente creado para YOLOv8. Este conjunto de datos se compone de imágenes variadas que representan diferentes tipos de vehículos, como automóviles, motos, buses, camiones y bicicletas. El dataset se divide en tres segmentos principales: uno para entrenamiento, con 1197 imágenes; otro para validación, que incluye 117 imágenes; y un último para pruebas, con 16 imágenes. A pesar de que el dataset fue inicialmente recopilado para YOLOv8, se realizó una adaptación cuidadosa para que pudiera ser empleado eficazmente en el entrenamiento de un modelo basado en YOLOv5, una versión previa del algoritmo. Este proceso de adaptación fue crucial para asegurar la compatibilidad y eficacia del modelo en la tarea de detección de vehículos.

Tecnologías

Para el espacio de trabajo en que se desarrolló la aplicación, se utilizaron los ambientes de Google Colab contando con el plan “Pay As You Go”, el cual brinda una cantidad de 100 unidades informáticas para la GPU T4. El desarrollo del sistema se estableció usando Python 3.9, junto con el framework de Flask para el servidor web donde se hospeda el modelo de machine learning.

El sistema de reconocimiento automático de matrículas vehiculares desarrollado emplea una serie de fases trascendentales recomendadas por distintos autores para asegurar un funcionamiento eficiente en tiempo real, las cuales se encuentran representadas en la Fig. 2. Estas incluyen la extracción de placas vehiculares destacada en [11], la detección de zonas de interés (ROI) enfatizada en [15] y el reconocimiento óptico de caracteres (OCR), fase en la que en [23] se sugieren métodos avanzados para enfrentar condiciones de iluminación diversas.



Fig. 2. Flujo desarrollado (Elaboración propia).

Estas fases constituyen el eje central del sistema ALPR y son el reflejo de una integración de conocimientos y técnicas recomendadas por diversos autores con énfasis en la materia. Se logró implementar un flujo de trabajo que comienza con la detección del área de interés en la que se presente un vehículo de forma precisa, siguiendo las recomendaciones en [22], previa a la identificación y captura de la matrícula del vehículo. Para ambos modelos de detección de objetos se decidió utilizar el algoritmo YOLO, principalmente para seguir la recomendación en [16] y [17], sobre enfocar los esfuerzos en métodos que utilizan redes neuronales convolucionales. Para el desarrollo del modelo, se utilizaron las versiones del algoritmo v5 y v8. Debido al tamaño de los datasets empleados, se eligió el tamaño de modelo “S”. Los modelos YOLO poseen diferentes letras que representan los tamaños de los modelos preentrenados. Estos modelos varían en complejidad y capacidad de detección de objetos, con tamaños que van desde “n” (más ligero) hasta “x” (más pesado), lo que afecta la precisión y el rendimiento del modelo en tareas de detección de objetos, para la ejecución de pruebas.

Este proceso se optimiza mediante un preprocesamiento que incluye técnicas de binarización, dilatación, erosión y desenfoque gaussiano, los cuales son de suma importancia para preparar las imágenes para el proceso de OCR. El sistema utiliza EasyOCR, una librería de código abierto y robusta, para la detección de textos en distintas zonas de una imagen. Este enfoque permite una lectura precisa de las matrículas en condiciones variables y complejas.

En el siguiente apartado se detallan los resultados obtenidos usando estas bases y tecnologías para el entrenamiento del modelo de visión computacional y cómo se desenvuelven bajo los criterios establecidos.

IV. ANÁLISIS DE RESULTADOS

A continuación, se describen los resultados obtenidos usando la metodología planteada sobre el conjunto de tecnologías de desarrollo detallado en la Fig.3 y con el conjunto de datos descrito con anterioridad.



Fig. 3. Stack empleado para el desarrollo del sistema ALPR (Elaboración propia).

En primer lugar, se contempló el uso de modelos de la familia de EfficientNet, pero fueron descartados al estar orientados principalmente al área de clasificación de objetos, en lugar de la detección de objetos en tiempo real, en contraste con los modelos de la familia

Se empleó YOLO, que posee una gran capacidad para realizar detecciones en tiempo real, utilizando la menor cantidad de recursos computacionales y priorizando la eficiencia. Ambos modelos, YOLOv5 y YOLOv8, fueron entrenados durante el mismo número de épocas y tomando el mismo conjunto de datos internacionales descrito en el apartado de métodos.

Tabla 2. Parámetros y resultados de entrenamiento con los modelos de la familia YOLO para detección de placas (Elaboración propia).

Modelo	YOLOv5	YOLOv8
Tamaño	S	S
Épocas	100	100
background vs. license_plate (Data de pruebas)	96 %	92 %
Confianza F1	0,96/0,442	0,96/0,524
Precision-Confidence	1,00/0,821	1,00/0,558
Precision-Recall	0,966 mAP@0,5	0,965 mAP@0,5
Recall-Confidence Curve	0,98/0,000	0,98/0,000
Modelo	YOLOv5	YOLOv8

De acuerdo con la Tabla 2, el modelo v5 exhibió una leve superioridad en la tarea de distinguir entre el fondo de la imagen y las matrículas, de modo que se alcanzó una precisión del 96 %, en comparación con el 92 % del modelo v8. El rendimiento de ambos tuvo una puntuación de 0,960, lo que indica alta precisión y sensibilidad en la detección de matrículas. Sin embargo, v8 tuvo un ligero incremento en términos de confianza, con un puntaje de 0,524 frente al 0,442 de v5. Este último mostró una mayor confianza (0,821) en comparación con YOLOv8 (0,558), lo que sugiere que mientras YOLOv5 fue más confiable, YOLOv8, aunque es preciso, fue menos consistente en sus predicciones.

Así, dichos resultados se representan también por medio de una matriz de confusión en la Fig.4 y Fig. 5, correspondientes al modelo YOLOv5 y YOLOv8, respectivamente.

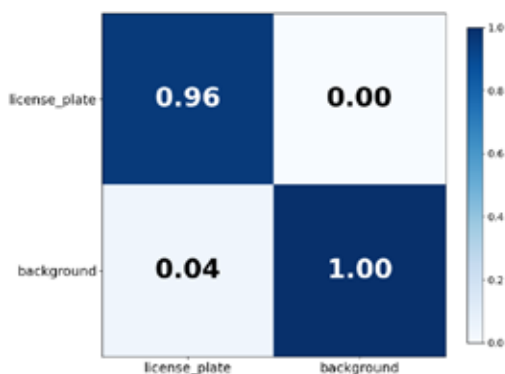


Fig. 4. Matriz de confusión matrículas vehiculares con YOLOv5 (Elaboración propia).

La matriz de confusión en la Fig. 4 evalúa el rendimiento del modelo en la clasificación de matrículas vehiculares. De acuerdo con esta, se proyectan los siguientes resultados sobre el modelo YOLOv5 comparando fondos y placas en la imagen:

- Verdaderos positivos: Las matrículas se clasificaron correctamente el 96 % del tiempo.
- Falsos positivos: A veces se clasificó erróneamente algún elemento del fondo como matrículas en un 4 % de los casos.
- Falsos negativos: No se perdieron matrículas.
- Verdaderos negativos: Todo el fondo se clasificó correctamente.

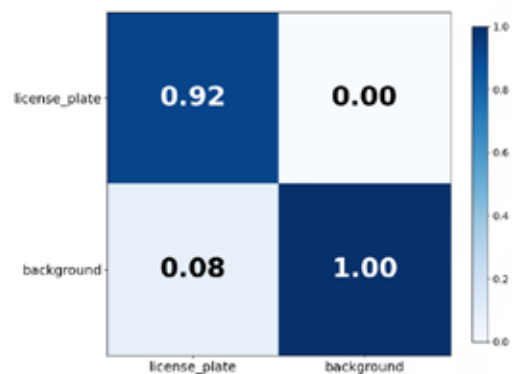


Fig. 5. Matriz de confusión matrículas vehiculares con YOLOv8 (Elaboración propia).

La Fig. 5 muestra una matriz de confusión para la detección de matrículas vehiculares utilizando el modelo YOLOv8. Los resultados bajo las mismas condiciones de la Fig. 4 son los siguientes:

- Verdaderos positivos: Las matrículas se clasificaron correctamente el 92 % del tiempo.
- Falsos positivos: A veces se clasificó erróneamente algún elemento del fondo como matrículas en un 8 % de los casos.
- Falsos negativos: No se perdieron matrículas.
- Verdaderos negativos: Todo el fondo se clasificó correctamente.

Dado que es evidente la consistencia altamente precisa del modelo YOLOv5, con base en la Tabla 2 y la Fig. 4, este fue seleccionado para conformar parte del flujo sistema en su rol de detección de placas vehiculares. La salida de este genera una o varias zonas de interés en la que se encuentra la placa vehicular identificada dentro de la imagen, dichas salidas se identifican en la Fig.6, donde se detectan las 2 placas presentes en la imagen de entrada.



Fig. 6. ROI generadas por el modelo de detección de placas (Elaboración propia).

Sin embargo, durante las etapas de pruebas y validación con el modelo de placas, se identificó un punto problemático a largo plazo, pues ciertos elementos viales que corresponden con características de placas internacionales levantan una detección falsa. Un ejemplo de esto se puede observar en la Fig. 7 en la que una señalización de tránsito en forma de rectángulo con la frase “4-WAY” es identificada como una placa vehicular.



Fig. 7. Detección de una señal de tránsito como una placa vehicular (Elaboración propia).

Esto se debe principalmente a que el dataset utilizado en el entrenamiento posee placas vehiculares internacionales, principalmente de países como Reino Unido y la India. Dichos países poseen matrículas con características menos diversas, pero con mayor cantidad de formatos, contrario a las placas de Costa Rica, las cuales son fácilmente identificables, debido a sus características propias y poseer variaciones dentro de las mismas. En consecuencia, ciertas señalizaciones de tránsito, carteles o letreros se tradujeron en un falso positivo, pues cumplen con las características de las placas de los vehículos presentes en el conjunto de datos de entrenamiento, siendo estas:

La presencia de un cuadrado o figura rectangular de algún color sólido.

- La presencia de letras y números.
- La presencia de un borde alrededor de las dos anteriores características.
- La presencia de figuras verticales, asemejándose a letras como “L”, “I”, “J”, o números como “7” o “1” encerradas en ángulos rectos.

Además, otros factores que influyeron en este tipo de detección son la cantidad de datos presentes durante el entrenamiento, las distintivas formas y colores de las placas internacionales. La confusión con letras del fondo que no se corresponden a zonas de interés en las que se ubique una placa vehicular, así como la muestra de marcas de agua, contribuyó también con la alerta de falsos positivos. Una muestra del dataset en la que se evidencian las características anteriormente descritas se representa en la Fig. 8.



Fig. 8. Muestra del dataset de entrenamiento utilizado (Elaboración propia).

Para abordar esta situación y como consecuencia de la falta de un dataset que permita realizar un entrenamiento exclusivamente basado en placas costarricenses, se decidió incluir en este punto un paso para centrar y dar mayor precisión a la búsqueda de placas vehiculares, lo que limita los datos que ingresan al modelo. Dicho proceso correspondió a la inclusión de un modelo de detección de vehículos como paso intermedio, como se sugirió en [22], esto permite primero ejecutar una búsqueda de zonas de interés con la presencia de vehículos, con ello, se reduce en gran medida la posibilidad de obtener detecciones no correspondientes a placas en el fondo. De igual forma, para este modelo de detección vehicular se optó por evaluar YOLOv5 y YOLOv8 igualmente en sus tamaños S. Los resultados de los mismos se detallan en la Tabla 3.

Tabla 3. Comparativa de modelos YOLOv5 y YOLOv8 para la detección de vehículos dentro de las imágenes (Elaboración propia).

Modelo	YOLOv5	YOLOv8
Tamaño	S	S
Épocas	10	10
Confianza F1	0,72/0,294	0,58/0,309
Precisión-Confidence	1,00/0,993	1,00/0,998
Precisión-Recall	0,771 mAP@0,5	0,642 mAP@0,5
Recall-Confidence Curve	0,90/0,000	0,91/0,000

De acuerdo con estos resultados, v5 superó a v8 con un valor de 0.72 en comparación con 0.58, indicando que YOLOv5 logró un mejor equilibrio entre precisión y recall en términos de confianza F1. La precisión-confianza fue alta en ambos modelos, lo que implica que ambos fueron muy precisos al realizar detecciones cuando tienen confianza en sus predicciones.

En la métrica de precisión-recall (mAP@0.5), YOLOv5 demostró un rendimiento superior con un valor de 0.771 en de contra un 0.642 de YOLOv8. Esto sugirió que v5 se desempeñó mejor en la tarea de detección de objetos, siendo más preciso y recuperando más objetos relevantes. La curva de recall-confianza muestra que ambos modelos tuvieron valores de recall cercanos al 90 % incluso cuando la confianza fue baja (0,000), indicando que ambos modelos fueron capaces de identificar objetos con una alta tasa de recall incluso en situaciones de baja confianza.

De igual forma, los modelos YOLOv5 y YOLOv8 se representa en la Fig. 9 y Fig. 10, mediante una matriz de confusión que contempla las etiquetas presentes en el dataset vehicular usado durante el entrenamiento.

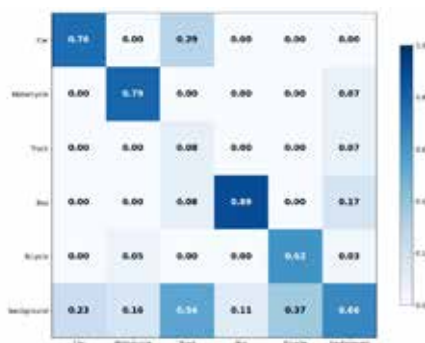


Fig. 9. Matriz de confusión vehículos con YOLOv5 (Elaboración propia).

La Fig. 9 presenta la matriz de confusión para la detección de vehículos utilizando el modelo YOLOv5, de la cual se obtuvieron los siguientes resultados: los automóviles se clasificaron correctamente el 76 % de las veces, las motocicletas el 79 %, mientras que los camiones mostraron una precisión más baja con solo un 37 % de predicciones acertadas. Por otro lado, los autobuses y las bicicletas presentaron altas precisiones, con un 89 % y 87 %, respectivamente. Sin embargo, se observó que el fondo se clasifica erróneamente como camiones en un 54 % de los casos.

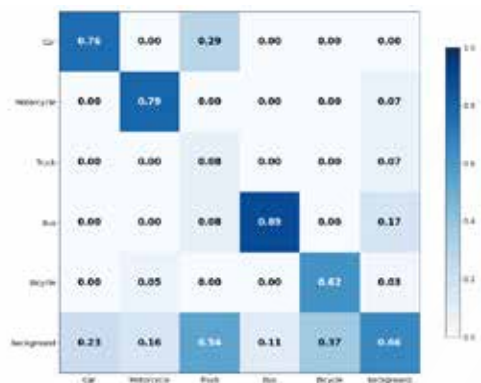


Fig. 10. Matriz de confusión vehículos con YOLOv8 (Elaboración propia).

La Fig. 10 muestra la matriz de confusión generada por el modelo YOLOv8 para la detección de vehículos. Según los resultados, los automóviles se clasificaron correctamente en un 76 % de los casos, mientras que las motocicletas alcanzaron una precisión del 79 %. Los autobuses presentaron una alta precisión con un 89 %, y las bicicletas también tuvieron un desempeño destacable con un 62 %. Sin embargo, los camiones registraron una precisión considerablemente baja, con solo un 8 % de predicciones correctas. Además, se observó que el fondo se clasificó erróneamente como camiones en un 54 % de las ocasiones.

Tomando como base ambas matrices en la Fig. 9 y la Fig. 10, y los resultados obtenidos en la Tabla 3, YOLOv5 destacó en comparación con YOLOv8 al presentar una efectividad superior para la tarea de detección de objetos usando el mismo conjunto de datos, cubriendo también una mayor precisión con variedad de clases. Bajo estos criterios, terminó siendo seleccionado finalmente para formar parte del flujo de detecciones.

Es relevante destacar que los resultados obtenidos con los modelos YOLOv5 y YOLOv8, tanto para la detección de placas como para la detección vehicular, corresponden a hallazgos experimentales obtenidos bajo las condiciones específicas del

estudio, las cuales están detalladas en las Tablas 2 y 3. No se ha realizado un análisis exhaustivo de las diferencias estructurales entre YOLOv5 y YOLOv8 que explique por qué YOLOv5 mostró un mejor desempeño en estos casos específicos, más allá de las métricas de rendimiento observadas. Los resultados obtenidos fueron los factores determinantes para la selección del modelo más adecuado para las tareas específicas de este proyecto.

Con la implementación del modelo de detección vehicular seleccionado, las salidas del este se representan en la Figura 11, y corresponden a ROI con vehículos presentes, dichas zonas de interés se ingresan como entrada para el modelo de detección de placas vehiculares. Al utilizar estos datos, se evita alimentar con elementos que no sean vehículos, lo cual limita el ingreso de elementos similares a placas internacionales.



Fig. 11. Salidas del modelo de detección vehicular (Elaboración propia).

Adicionalmente, se limitó la salida del modelo para solo devolver ROI con vehículos con un ancho o altura mayor a 100px, esto con el fin de evitar que el modelo de placas analice vehículos muy lejanos cuya extracción de placa sea improbable dada su resolución. Con este enfoque, la corrección del falso positivo planteado anteriormente se demuestra en la Fig. 12, ignorando la señal de tránsito mal identificada previamente.



Fig. 12. Detección de una señal de tránsito como una placa vehicular corregida (Elaboración propia).

Mediante la librería OpenCV se recortaron las ROIs de las placas, ingresando una entrada mucho más limpia en el flujo de OCR. Siguiendo la tendencia provista por los autores referentes a sistemas ALPR, el uso de un OCR para la extracción de los caracteres alfanuméricos en las placas extraídas se presenta como una de las mejores y más eficientes opciones con este objetivo.

Inicialmente, se planteó agregar Pytesseract, una implementación en Python de Tesseract desarrollada por Google, como parte del flujo con OpenCV, igualmente se consideró KerasOCR, una biblioteca propia de Keras. Sin embargo, ambas herramientas se orientaban principalmente a la identificación de oraciones y párrafos de un texto. Por este motivo, se recurrió a EasyOCR, con el cual se aplican las funciones de preprocesamiento y se alimenta al OCR con la última imagen obtenida del flujo de detección y extracción. Previo a la lectura de caracteres, se transforma la imagen a escala de grises para descartar el uso de colores que puedan generar ruido en la entrada. Luego de esto, se aplicó, mediante el método Otsu, una operación de umbralización binaria sobre la imagen para obtener una versión binaria, agrupando los píxeles de la misma en blancos y negros. Después de esto, se aplican dos funciones con una matriz kernel 3x3. Primero, se realiza una dilatación de imagen para unir regiones de los caracteres que puedan estar separadas, posteriormente se hace una erosión de la imagen para eliminar ruido sobrante y refinar los caracteres. Por último, se aplica un desenfoque gaussiano para suavizar los bordes de los caracteres. El flujo consecuente de filtros se representa en la Fig. 13.



Fig. 13. Transformaciones y filtros aplicados a las placas (Elaboración propia).

Con el fin de probar la pertinencia del modelo de detección de placas al exponerse a matrículas vehiculares de Costa Rica, se recopiló un dataset compuesto por imágenes con vehículos costarricenses por medio de web scraping. Las imágenes fueron extraídas de sitios de compra y venta de vehículos. La cantidad de imágenes recopiladas fue de 638 imágenes, de las cuales se etiquetaron 729 placas en formato PASCAL VOC. Una muestra se evidencia en la Fig. 14.

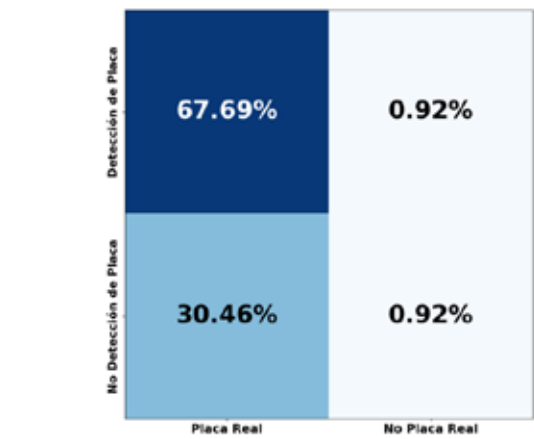


Fig. 14. Muestra de imágenes de vehículos costarricenses recopiladas mediante web scraping (Elaboración propia).

La Fig. 15 muestra la matriz de confusión del sistema de detección de placas vehiculares generada al realizar detecciones sobre el dataset de pruebas previamente descrito. Los resultados mostraron un 67,69 % de verdaderos positivos de placas cuando estas están presentes frente a un 30,46 % de falsos negativos, lo que indica que en estos casos el sistema no detectó la placa a pesar de que estaba presente. Sin embargo, los falsos positivos, es decir, las predicciones incorrectas de detección de placa cuando no había una placa real, representan solo un 0,92 %, al igual que los verdaderos negativos, donde el sistema acertó al no detectar una placa cuando efectivamente no la había. En este sentido, los resultados destacan la necesidad de mejorar la precisión del sistema, especialmente en la reducción de los falsos negativos.

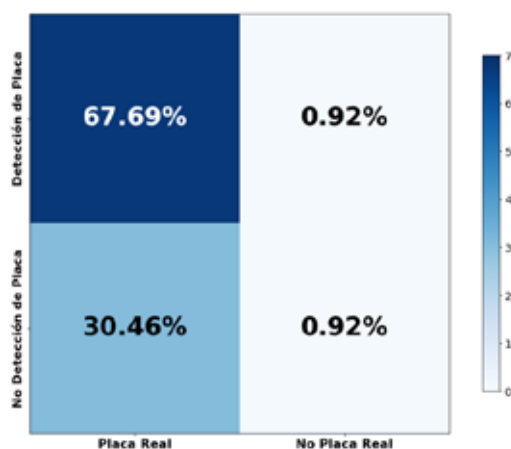


Fig. 15. Matriz de confusión porcentual del desempeño del modelo con el dataset recopilado (Elaboración propia).

Para ilustrar esto, algunos de los ejemplos de detección más importantes, junto con sus porcentajes de confianza, se agregaron en la Fig.16.



Fig. 16. Resultados relevantes durante la prueba de detección (Elaboración propia).

En la imagen los casos 1, 4, 5, 6, 7, 8, 9 y 10 se corresponden a detecciones de placas completas en distintos ángulos, proximidad a la cámara o condiciones de luz. Por otro lado, las detecciones 2 y 3 representan placas que el modelo logró identificar a pesar de que están incompletas, tapadas o cortadas en la imagen. Por el contrario, las detecciones 11 y 12 corresponden a falsos positivos. Particularmente en el caso 11, este corresponde al logo de la marca del vehículo, contando con una estructura similar a la de placas internacionales presentes en el dataset de entrenamiento mostrado en la Fig. 8. Por otro lado, en el caso 12, la detección se dio en la esquina de la imagen, dado que se forma un ángulo de 90 grados y se cierra una forma cuadrada debido a la estructura de la cerca, da como resultado una semejanza a una placa con caracteres rectos como el número 1 o 7, lo que en consecuencia puede estar causando una detección errónea.

V. DISCUSIÓN Y CONCLUSIONES

El desarrollo del sistema ALPR descrito en este artículo representa un avance en cuanto a la seguridad vial en Costa Rica, dado que es posible reconocer placas vehiculares nacionales en entornos urbanos. Este enfoque es útil, puesto que permite identificar automáticamente las placas de los vehículos en imágenes o vídeos, lo que resulta beneficioso para la seguridad, el control de acceso y la gestión del tráfico.

No obstante, si bien es cierto que este método es funcional con placas costarricenses, los casos 11 y 12 presentes en la Fig. 15 evidencian que existen diversas limitaciones que abren la puerta a futuras mejoras, especialmente en lo que respecta a la precisión y eficiencia del sistema. Las principales limitaciones actuales del sistema se resumen en:

Impacto del dataset internacional en la detección de placas

El uso de un dataset con placas internacionales (principalmente de Reino Unido e India) generó la problemática en que elementos viales, señalizaciones de tránsito, carteles o letreros con características similares a las placas internacionales (forma rectangular, presencia de letras/números, borde, figuras verticales) a menudo fueron clasificados erróneamente como placas vehiculares, resultando en falsos positivos.

Necesidad de un modelo de detección vehicular auxiliar

Debido a los falsos positivos generados por el modelo de detección de placas al confundir elementos del fondo con placas internacionales, se incorporó un modelo de detección vehicular que limita el modelo de placas a solo las áreas donde se detectan vehículos. Si bien esto mejora la precisión, agrega un paso adicional al flujo de procesamiento, lo que podría impactar la eficiencia en ciertos escenarios.

Precisión mejorable con placas costarricenses:

Al probar el modelo de detección de placas entrenado con el dataset internacional con un dataset de pruebas recopilado con datos costarricenses, se obtuvieron resultados que, aunque funcionales, mostraron un margen de mejora. Específicamente, se registró un 67,69 % de verdaderos positivos y un 30,46 % de falsos negativos, lo que significa que en una proporción considerable de casos, el sistema no detectó la placa a pesar de estar presente.

Limitaciones en la corrección de perspectiva y ejes

La corrección de los ángulos de las placas para estandarizar las imágenes y facilitar la lectura mediante OCR no pudo implementarse debido a limitaciones técnicas de las herramientas utilizadas. Este proceso sería clave para mejorar la precisión en la lectura de placas en diferentes perspectivas.

Estas limitaciones destacan áreas clave donde el sistema actual enfrenta desafíos y donde se concentran las recomendaciones para futuros trabajos. Los hallazgos descritos y las limitaciones descritas con anterioridad tienen varias implicaciones importantes para futuras investigaciones en el área de reconocimiento de placas vehiculares en Costa Rica:

Necesidad de un dataset local exclusivo

La principal recomendación para futuras investigaciones es la creación de un dataset exclusivo para Costa Rica. El uso de un dataset local permitiría mejorar considerablemente el rendimiento y la precisión del modelo, eliminando los falsos positivos generados por placas extranjeras o elementos similares a estas. Además, un dataset costarricense adecuado podría eliminar la necesidad del paso intermedio de detección vehicular, lo que agilizaría el procesamiento de imágenes en tiempo real. Las placas costarricenses tienen características propias suficientes para distinguirlas directamente del fondo.

Validación de modelos en contextos locales

Aunque modelos avanzados como YOLOv5 y YOLOv8 han mostrado un buen desempeño en datasets internacionales, es fundamental validarlos en contextos locales. Comparar los resultados obtenidos con modelos entrenados en datos costarricenses permitirá determinar la efectividad de estos modelos y ajustar su rendimiento a las características locales.

Enfoque en optimización del flujo

Futuras iteraciones del trabajo deben centrarse en optimizar el flujo descrito o explorar arquitecturas de extremo a extremo que no requieran pasos intermedios, especialmente si se dispone de un dataset local adecuado. Esto no solo mejoraría la precisión, sino también la eficiencia del procesamiento.

Áreas de mejora técnica

Se identificaron varias áreas técnicas para mejorar, como la corrección de la perspectiva y los ejes de las placas, lo que facilitaría el preprocesamiento y mejoraría la calidad de los datos para el OCR. Aunque esta mejora no fue posible en este estudio, sigue siendo de suma importancia para futuras investigaciones.

Uso de recursos computacionales

El uso de GPU más avanzadas en futuras iteraciones del sistema podría facilitar el entrenamiento de modelos más complejos y la incorporación de datasets más grandes, eliminando la potencia computacional como un factor limitante. Esta mejora permitiría escalar el sistema y obtener mejores resultados con modelos más robustos.

RECOMENDACIONES

En la búsqueda de alcanzar un escenario técnico óptimo, los autores poseen interés en expandir el dataset a un conjunto de datos exclusivamente costarricenses. Todo esto con el objetivo de brindar continuidad y extender el trabajo desarrollado.

Finalmente, se considera la reiteración y la mejora continua de las distintas etapas del sistema como un proceso fundamental para el avance y la optimización del proyecto. Se invita a otros investigadores a participar en este campo, superando los resultados actuales y explorando nuevas posibilidades y aplicaciones del sistema ALPR propuesto, dando un énfasis especial al análisis de la relación entre sistemas ALPR y la expansión potencial de ciudades inteligentes en Costa Rica.

AGRADECIMIENTOS Y/O FINANCIAMIENTO

Este trabajo contó con el apoyo de la Universidad Cenfotec, patrocinando la investigación y formación de los autores dentro del marco de becas universitarias de excelencia. También se extiende un agradecimiento especialmente al doctor Robert Beck y al doctor Ulises Agüero.

REFERENCIAS

- [1] D. Solano Cambroner, “Anuario estadístico de accidentes de tránsito con víctimas en Costa Rica 2024,” Consejo de Seguridad Vial, 2024. [En línea]. Disponible en: <https://www.csv.go.cr/estad%C3%ADsticas>
- [2] Organismo de Investigación Judicial (OIJ), “El robo de vehículos se incrementa y usted puede salvar el suyo,” Sitio Oficial del Organismo de Investigación Judicial. [En línea]. Disponible en: <https://sitiooij.poder-judicial.go.cr/index.php/comunicacion/noticias/boletin-oij-actualidad-policial/item/22329-el-robo-de-vehiculos-se-incrementa-y-usted-puede-salvar-el-suyo>
- [3] L. N. Zeledón, “Investigación en Informática: el enfoque alternativo,” *Technology Inside by CPIC*, vol. 5, pp. 1–15, 2020.
- [4] J. Biolchini, P. G. Mian, A. C. C. Natali y G. H. Travassos, “Systematic review in software engineering,” *System Engineering and Computer Science Department COPPE/UFRJ, Tech. Rep. ES-679.05*, 2005.
- [5] P. Kulkarni, A. Khandebharad, D. Khope y P. U. Chavan, “License plate recognition: a review,” 2012 Fourth International Conference on Advanced Computing (ICoAC), pp. 1–8, dic. 2012.
- [6] S. Du, M. Ibrahim, M. Shehata y W. Badawy, “Automatic license plate recognition (ALPR): A state-of-the-art review,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 23, no. 2, pp. 311–325, 2012.
- [7] T. Björklund, A. Fiandrotti, M. Annarumma, G. Francini y E. Magli, “Robust license plate recognition using neural networks trained on synthetic images,” *Pattern Recognition*, vol. 93, pp. 134–146, 2019.
- [8] W. Wang, J. Yang, M. Chen y P. Wang, “A light CNN for end-to-end car license plates detection and recognition,” *IEEE Access*, vol. 7, pp. 173875–173883, 2019.

- [9] R. Laroca, E. Severo, L. A. Zanlorensi, L. S. Oliveira, G. R. Gonçalves, W. R. Schwartz y D. Menotti, "A robust real-time automatic license plate recognition based on the YOLO detector," 2018 International Joint Conference on Neural Networks (IJCNN), pp. 1–10, jul. 2018.
- [10] B. Y. Amirgaliyev, C. A. Kenshimov, K. K. Kuvatov, M. Z. Kairanbay, Z. Y. Baibatyr y A. K. Jantassov, "License plate verification method for automatic license plate recognition systems," 2015 Twelfth International Conference on Electronics Computer and Computation (ICECCO), pp. 1–3, sep. 2015.
- [11] L. Xu, W. Shang, W. Lin y W. Huang, "License plate detection methods based on OpenCV," 2021 21st ACIS International Winter Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD-Winter), pp. 11–16, ene. 2021.
- [12] S. Yu, B. Li, Q. Zhang, C. Liu y M. Q. H. Meng, "A novel license plate location method based on wavelet transform and EMD analysis," *Pattern Recognition*, vol. 48, no. 1, pp. 114–125, 2015.
- [13] H. Lin, J. Zhao, S. Li y G. Qiu, "License plate location method based on edge detection and mathematical morphology," 2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), vol. 1, pp. 853–857, jun. 2020.
- [14] I. Giannoukos, C. N. Anagnostopoulos, V. Loumos y E. Kayafas, "Operator context scanning to support high segmentation rates for real time license plate recognition," *Pattern Recognition*, vol. 43, no. 11, pp. 3866–3878, 2010.
- [15] C. S. Ahn, B. G. Lee, S. S. Yang y S. C. Park, "Design of car license plate area detection algorithm for enhanced recognition plate," 2017 4th International Conference on Computer Applications and Information Processing Technology (CAIPT), pp. 1–4, ago. 2017.
- [16] M. A. H. Bakhtan, M. Abdullah y A. Abd Rahman, "A review on license plate recognition system algorithms," 2016 International Conference on Information and Communication Technology (ICICTM), pp. 84–89, may. 2016.
- [17] M. F. Sadique y S. R. Haque, "A comparative study of license plate detection and recognition techniques," 2019 22nd International Conference on Computer and Information Technology (ICCIT), pp. 1–6, dic. 2019.
- [18] I. V. Pustokhina, D. A. Pustokhin, J. J. Rodrigues, D. Gupta, A. Khanna, K. Shankar et al., "Automatic vehicle license plate recognition using optimal K-means with convolutional neural network for intelligent transportation systems," *IEEE Access*, vol. 8, pp. 92907–92917, 2020.
- [19] D. Pu, N. Gu y X. Zhang, "A robust and real-time approach for license plate detection," 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), pp. 1–7, jul. 2018.
- [20] A. Menon y B. Omman, "Detection and recognition of multiple license plate from still images," 2018 International Conference on Circuits and Systems in Digital Enterprise Technology (ICCSDET), pp. 1–5, dic. 2018.
- [21] B. S. Shobha y R. Deepu, "A review on video based vehicle detection, recognition and tracking," 2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS), pp. 183–186, dic. 2018.
- [22] H. C. Kuo, J. S. Lee y S. T. Kao, "An autonomous license plate detection method," 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, pp. 110–113, sep. 2009.
- [23] A. P. Palanivel, T. Vigneshwaran, M. S. Arappadhan y R. Madhanraj, "Automatic number plate detection in vehicles using faster R-CNN," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), pp. 1–6, jul. 2020.
- [24] A. I. Mora Vargas, "Guía para elaborar una propuesta de investigación," *Revista Educación*, vol. 29, no. 2, pp. 77–97, 2012. [En línea]. DOI: <https://doi.org/10.15517/revedu.v29i2.2241>



TIME AND COST OF QUALITY IN SOFTWARE VERIFICATION: A COMPARATIVE SURVEY OF RUP, EVOLUTIONARY PROTOTYPING, AND SCRUM

MSI. Esteban Sanabria-Mora

*Master's Degree in Computer Science and Information Systems
Instituto Tecnológico de Costa Rica. San José, Costa Rica
esteban.sanabria@tec.ac.cr*

ABSTRACT

Quality assurance has become a fundamental component of modern software development, as project success is increasingly linked to the ability to deliver reliable and high-quality software within defined time and budget constraints. As a result, understanding how time and cost are affected by quality-related activities has become a major concern for organizations. This paper analyzes the role of the Time of Quality (ToQ) and the Cost of Quality (CoQ) within the software verification process of projects developed using three different methodologies: the Rational Unified Process (RUP), evolutionary prototyping, and Scrum. Through a comparative analysis of how each methodology integrates testing and verification activities throughout the software development lifecycle, the study highlights differences in the estimation, management, and impact of quality-related time and cost. The results provide a conceptual reference to support the selection of a software development methodology based on its ability to manage quality assurance efforts while balancing time and budget constraints.

Keywords: Software development methodology, time, cost, quality assurance, verification, project management, prototyping, Scrum, RUP.

I. INTRODUCTION

As with all products and services in industry, software applications require systematic quality assurance (QA) processes to ensure correctness, completeness, and compliance with customer and business requirements [1]. Beyond validating the final product, QA also supports the control and improvement of the software engineering process throughout the project lifecycle. Consequently, quality assurance constitutes a broad and complex discipline that directly influences both project outcomes and resource consumption.

Quality assurance activities are commonly grouped into verification, prevention, and correction of nonconformities [2]. Among these dimensions, software verification plays a particularly critical role, as it involves evaluating software artifacts at different stages of development through reviews, inspections, and testing activities. Verification tasks are directly associated with measurable project variables such as time and cost, since they require dedicated effort, tooling, and human resources. Given the wide scope of QA, this paper focuses specifically on the verification phase, which allows a clearer analysis of how quality-related time and cost are managed within different development methodologies.

The aim of this study is to compare how three software development methodologies (the Rational Unified Process, evolutionary prototyping, and Scrum) address software verification activities and how these approaches affect the Time of Quality (ToQ) and the Cost of Quality (CoQ). The comparison is based on methodological characteristics defined in each model, together with commonly reported practices in the software industry regarding testing effort, iterative verification, and quality-related cost distribution. The findings provide a conceptual reference for organizations seeking to select a software development methodology that aligns with their quality assurance objectives and their constraints in terms of time and cost.

A. Objective

To examine and compare the time and cost required to ensure software quality at the verification level in projects developed using RUP, evolutionary prototyping, and Scrum.

B. Problem and Justification

Software development methodologies continue to evolve in response to technological advances and changing organizational needs. As a result, a methodology that is effective in one context may be less suitable in another. This highlights the need not only to analyze software development models, but also to examine how they incorporate and manage verification activities related to quality assurance.

In addition, despite the central role of quality assurance in software projects, many organizations often lack clear guidance on how different development methodologies affect the time and cost associated specifically with verification activities. For these reasons, this study justifies a comparative analysis of RUP, evolutionary prototyping, and Scrum, focusing specifically on how each methodology addresses verification activities and how these practices impact the Time of Quality and the Cost of Quality. The results aim to provide a conceptual reference to support informed decision-making when selecting a software development methodology aligned with quality, time, and budget objectives.

II- BACKGROUND AND RELATED WORK

Nowadays, software development models have become widespread across most companies. In this context, selecting an appropriate software model that enables the early detection and reduction of errors during the quality assurance (QA) phase has become increasingly important. Some organizations choose to create their own models to meet specific needs, while others implement existing software models with specific adaptations. Regardless of the approach, framework, or software model adopted, ensuring the quality of software applications has become a critical concern, as have become more demanding than ever in terms of software reliability and performance. Some traditional software models, such as the Waterfall model, which are still used despite the rise of agile methodologies, treat software verification as a phase separate from other processes, such as analysis, design, and development.

In other words, functionality verification of functionalities, or testing, occurs during a single phase and only after development artifacts have been completed. By contrast, companies such as IBM, which have adopted the Rational Unified Process (RUP) as their primary software development model, state that RUP includes a dedicated testing phase. However, in RUP, this phase occurs

at the end of each iteration that introduces new functionality to the increment—that is, the set of implemented features—rather than solely at the end of the project. To illustrate the testing phase within RUP, Fig. 1 presents this phase in comparison with the other phases that comprise the development lifecycle of a software project.

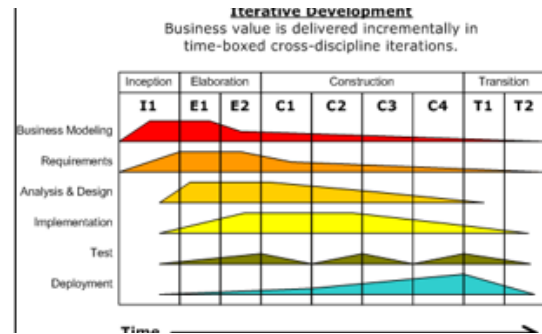


Fig. 1. Development phases in RUP [3].

Unlike traditional models, agile methodologies view quality assurance not as a standalone activity but as an integral part of development. For instance, in evolutionary prototyping, development and testing activities are tightly interwoven. Developers release functional prototypes as soon as they are testable, allowing users to provide feedback while development continues. Scrum manages quality assurance differently from both RUP and evolutionary prototyping. Each user story includes acceptance criteria that define the conditions under which the story is considered complete. Additionally, all sprints are governed by the Definition of Done, a set of quality criteria used to verify the correctness, completeness, and overall quality of the increment produced during a sprint.

Overall, quality verification is handled differently depending on the development model and organizational practices. Agile methodologies incorporate verification iteratively, while traditional models tend to separate verification from development activities, either iteratively, as in RUP, or sequentially, as in the Waterfall model.

III- THEORETICAL FRAMEWORK

A. Concepts

1) *Quality Assurance*

According to Wiley and Wiley, the concept of quality assurance encompasses all handled of the software development lifecycle, ensuring that its implementation complies with both organizational and client quality standards, and that potential inconsistencies or errors are identified and corrected [4]. This concept differs from software testing, which focuses specifically on evaluating the software product by assessing the correctness of the results produced when executing operations that users are expected to perform once the product is released. In more colloquial terms, as stated by Joseph Man, “testing should intentionally attempt to make things go wrong to determine if things happen when they shouldn’t, or things don’t happen when they should” [4].

2) *Verification testing*

The verification phase within quality assurance refers to the practice of tracking and evaluating all software development activities to certify their compliance with expected behavior. This phase is not limited solely to programming tasks only but also encompasses all other project-related activities that contribute to the software product as the primary outcome. Historically, the verification phase of QA was often conceived as a process applied only at the end of the software development lifecycle. However, it is now widely recognized that verification should take place throughout all phases of software creation, including requirements gathering, analysis, design, programming, implementation, and formal user training.

3) *The time of quality (ToQ)*

The term “time of quality” refers to the number of working hours devoted to detecting and correcting defects identified through the use of testing strategies and techniques during the verification process. The time required to implement quality assurance practices in the verification phase depends largely on the complexity and scope of the application under test. For example, when automated testing is employed instead of manual testing, the time needed to assure software quality can be significantly reduced. Experts suggest that automated testing is particularly suitable for

large-scale applications with frequent releases. Conversely, when an application is small or not a suitable candidate for automated testing, manual verification becomes the preferred approach, which typically requires a greater investment of time.

4) *The cost of quality (CoQ)*

The term “cost of quality” refers to the costs incurred to correct defects that are discovered and addressed throughout the development and delivery of a software product. It may result from any of the following scenarios, according to the American Society for Quality [5]:

- Performance of unnecessary work because of errors, poor organization, or communication.

- Correction of defective results or errors.

- Analysis of the causes of the failures to prevent them from happening again.

Wiley & Wiley [4] state that the cost of applying quality assurance in software products often exceeds half the total cost of software development and maintenance. That is because “there are no fixed recipes for attacking the problem of verifying a software product as even the most experienced specialists do not have pre-cooked solutions but need to design a solution that suits the problem, the requirements, and the development environment” [4].

B. Software development methodologies for this survey

In order to understand the role of time and cost as key variables in quality assurance processes, it is essential to examine the software development methodologies through which software is produced. Accordingly, Scrum, the Rational Unified Process (RUP), and evolutionary prototyping are briefly introduced in the following sections.

1) *Scrum*

Scrum is a framework used to develop complex projects, which began in the early 1990s. It is mainly characterized by using an iterative and incremental method in the development and delivery of results, which is known as “sprint” and has an approximate duration of two weeks to one month. This framework was developed by Ken Schwaber, Jeff Sutherland and Mike Beedle, and is particularly suitable for projects with rapidly changing requirements [8]. The result of each sprint, which is composed of “user

stories (requirements in a format similar to Kanban visual cards), is an increment that is presented to customers on a regular basis for approval or feedback. In Scrum, there are four artifacts that promote a collaborative and transparent strategy of the work to be done, the work in progress, and the work done.

2) RUP

IBM defines “The Rational Unified Process” or RUP as “a generic software development methodology that works for varied types of software applications, organizations, levels of competitiveness, and size of projects. Its goal is to ensure the production of high-quality software that meets the needs of its end-users, within a predictable schedule and budget” [3]. RUP is an attempt to draw on the best features and characteristics of traditional software process models, such as Waterfall, and at the same time implement some of the best principles of agile software development [7]. RUP utilizes the Unified Modeling Language (UML) as its main tool to analyze and design software. This language allows developers, software architects, and other professionals involved in the creation of software products to clearly model requirements and bring them as close as possible to their implementation in a programming language.

3) Evolutionary prototyping

Evolutionary prototyping is a lifecycle model that enables software applications to be developed incrementally. This approach relies heavily on customer feedback, allowing modifications and the addition of functionalities at a pace determined by customer approval. The model is particularly useful when software requirements are complex, incomplete, or not clearly defined, as it helps users better understand and refine their needs throughout the development process.

This model operates through evolutionary releases of the application, in which the most critical requirements are prioritized, designed, and implemented to deliver functional components that users or customers can evaluate at any stage of the project. The model can be applied to a wide range of software development projects, as it provides timely and effective feedback from future users, contributing to continuous improvements in product correctness and quality.

IV- TIME AND COST OF QUALITY ASSURANCE IN THE THREE MODELS

Capgemini’s World Quality Report, published in 2015-16, found that today’s software development companies spend 35% of their budgets on QA and testing [9]. But is it the same case for all software models? This section addresses each model from the time and cost of QA’s perspective.

1) Time and cost of QA in RUP

In the Rational Unified Process (RUP), quality assurance activities are integrated across all project phases and involve all stakeholders involved in the assessment of the software product. Functional testing is not treated as a standalone activity performed by a separate team; rather, it is embedded within the development process itself [7].

The amount of time RUP devotes to verification testing depends on the number of features developed during the construction phase. For each iteration intended to deliver functional components to the customer, “all features are developed and integrated into the product, and all features are thoroughly tested” [3]. Although the construction phase is where testing and verification are most intensive, quality assessment is not limited to this stage. Verification activities may also occur during the inception, elaboration, and transition phases, even though new functionality is primarily introduced during the construction and transition stages. In this context, according to Pressman and Maxim [7], verification and testing activities typically consume between 30% and 40% of the total project effort in structured and iterative development models such as RUP.

From a cost perspective, RUP’s iterative verification strategy contributes to a more predictable and controlled Cost of Quality. By identifying defects early in the lifecycle, the process reduces the cost of rework, defect propagation, and late-stage corrections, which are typically more expensive. The structured nature of RUP allows organizations to allocate resources explicitly to quality-related activities within each iteration, making quality costs more traceable and manageable throughout the project. Consequently, RUP projects often exhibit a predictable and controlled Cost of Quality, typically representing close to one-third of the total project budget, which aligns with its focus on risk mitigation and high-quality deliverables.

2) Time and cost of QA in Scrum

The Scrum Body of Knowledge (SBOK) defines quality as the ability of a completed product or its deliverables to meet the acceptance criteria of all user stories, which is ultimately reflected in the achievement of the business value expected by the customer [8]. To ensure that a project meets its quality requirements, Scrum advocates maintaining an updated and prioritized product backlog and defining clear acceptance criteria for every user story. These criteria function as quality parameters against which each unit of functionality is evaluated. According to the SBOK, “clearly defined acceptance criteria are crucial for timely and effective delivery of the functionality defined in the user stories, which ultimately determines the success of the project in terms of time and cost” [8].

From a time perspective, Scrum manages the Time of Quality by integrating verification activities into every sprint. Because development proceeds incrementally, defects can be identified early through repeated testing after each iteration, rather than being discovered only at the end of the project. This iterative verification reduces the time required for extensive rework and minimizes delays associated with late-stage defect correction. Testing effort is commonly estimated as part of the sprint workload using story points [10].

From a cost perspective, Scrum distributes quality-related costs across sprints instead of concentrating them in a single testing phase. Quality assurance activities such as testing, documentation, and review are performed as part of regular sprint tasks, allowing teams to control the Cost of Quality incrementally. Although Scrum does not prescribe fixed percentages for quality-related effort, the use of story points, sprint planning, and acceptance criteria enables teams to estimate and adjust quality-related costs with reasonable accuracy throughout the project lifecycle. This continuous integration of quality practices increases the likelihood of achieving the desired level of quality while maintaining control over both time and budget.

3) Time and cost of QA in Evolutionary prototyping

Evolutionary prototyping is a software development model that enables faster delivery when compared to traditional models such as the Waterfall model. It is designed to leverage modern development tools that support rapid construction and modification of software artifacts. Rather than emphasizing extensive upfront planning, this model

prioritizes ongoing development, allowing requirements to evolve throughout the project based on continuous user feedback.

Development in evolutionary prototyping is conducted through short, time-boxed cycles that include implementation, testing, and verification activities. Each prototype represents an incremental version of the final product and is evaluated by users or customers to validate functionality, usability, and overall quality. Feedback gathered during these evaluations is incorporated into subsequent iterations, facilitating early defect detection and reducing the likelihood of costly rework at later stages of the project [11].

From a time perspective, the Time of Quality in evolutionary prototyping is closely intertwined with development activities, as verification occurs continuously alongside implementation. This integration can shorten feedback loops and accelerate defect identification. However, because verification tasks are not formally defined or standardized, the amount of time devoted to quality activities may vary significantly between iterations and projects, making precise estimation of ToQ challenging.

From a cost perspective, the Cost of Quality in evolutionary prototyping is similarly difficult to quantify. While early feedback and rapid validation may reduce the cost of late-stage defect correction, the absence of structured quality checkpoints, formal documentation, and predefined testing criteria can lead to inefficiencies, rework, and maintainability issues. These factors often result in unpredictable quality-related costs, particularly in projects with frequent requirement changes or poorly controlled iteration cycles.

Overall, evolutionary prototyping supports early verification and user-driven validation, which can positively impact both time and cost. However, compared to methodologies such as RUP and Scrum, the lack of explicit mechanisms for planning, measuring, and controlling quality activities limits its effectiveness in estimating the Time of Quality and Cost of Quality.

4) Comparison of the Time and cost of Quality in the three different software models

Table 1 summarizes the comparison of the Time of Quality (ToQ) and Cost of Quality (CoQ) specifically associated with the software verification phase in projects developed using RUP, Scrum, and evolutionary prototyping.

TABLE 1. Comparison of the Time and Cost of Quality in the three different software models.

Software Model/ Measure	ToQ	CoQ
RUP	Approximately one-third of the total project development time is devoted to verification and testing activities, distributed across all iterations, with a strong emphasis during the construction phase.	Approximately one-third of the total project budget is allocated to verification and testing activities, including test planning, execution, automation, and defect correction.
Scrum	Quality activities are continuously integrated into each sprint through acceptance criteria, Definition of Done, testing, refactoring, and continuous integration rather than being treated as a separate phase.	Quality-related costs are distributed across sprints and managed incrementally through sprint planning, acceptance criteria, continuous integration, and defect resolution activities.
Evolutionary prototyping	Time devoted to quality varies significantly across iterations and is often difficult to quantify due to evolving requirements, informal verification practices, and frequent user feedback cycles.	The cost of quality is highly variable and may increase over time due to frequent requirement changes, rework, limited architectural planning, and maintenance challenges in evolving prototypes.

As shown in Table 1, RUP and Scrum provide clearer mechanisms for estimating and managing the time and cost associated with the verification phase. At the same time, evolutionary prototyping lacks explicit structures for verification planning and measurement.

V- DISCUSSION

The comparison of the Time of Quality (ToQ) and Cost of Quality (CoQ) across the three software development models reveals substantial differences in how quality assurance activities are planned, executed, and controlled. The Rational Unified Process (RUP) exhibits the most structured approach, characterized by clearly defined phases and explicit verification activities. This structure enables relatively precise estimation of both the time and cost associated with quality assurance, making RUP particularly suitable for projects in which strict quality control, extensive documentation, and reliable budget forecasting are critical.

Scrum also integrates quality assurance throughout the development lifecycle but does so more flexibly and iteratively. Although its ToQ and CoQ are not expressed as fixed proportions, they can be reasonably estimated through sprint planning mechanisms, story point allocation, and clearly defined acceptance criteria and Definition of Done. This adaptability allows Scrum teams to continuously balance quality, time, and cost, making the methodology well-suited for projects with evolving requirements and frequent incremental deliveries.

In contrast, evolutionary prototyping lacks clearly defined metrics for ToQ and CoQ. While the model facilitates rapid development and early user feedback, the absence of formal quality checkpoints and structured planning can lead to inefficiencies that complicate time and cost estimation. Consequently, although evolutionary prototyping may be effective for exploratory or highly uncertain projects, it presents increased risks when rigorous control over quality-related time and cost is required.

Overall, methodologies with explicitly defined quality assurance practices, such as RUP and Scrum, offer stronger support for managing and estimating the time and cost associated with quality. Therefore, the selection of a software development model should consider not only development speed and flexibility, but also the organization's ability to plan, measure, and control quality-related activities within established time and budget constraints.

VI- CONCLUSIONS

Before initiating a software development project, it is essential to understand its characteristics and to clearly define how it will be executed. This preliminary analysis enables organizations to optimize resources, manage time and cost effectively, and ensure a balanced distribution of responsibilities and benefits among project participants.

An important finding of this study is that the selected development methodology directly influences team participation and workload distribution. Depending on the model adopted, certain roles may require greater involvement than others, which in turn affects both the effort invested and the final cost of the software product. Consequently, the choice of methodology has a direct impact on quality-related expenditures and timelines.

The findings of this paper confirm that time and cost are key variables in quality assurance processes and that their management is highly dependent on the software development methodology employed. The testing phase is shown to be a critical component of the software development lifecycle, with its execution varying according to the model and the methodological adaptations applied by each organization. Agile methodologies apply quality assurance iteratively at the verification level, while traditional models tend to separate verification from programming activities, either iteratively, as in RUP, or sequentially, as in the Waterfall model.

RUP exhibits the most structured and predictable approach, with explicitly defined verification activities integrated throughout the lifecycle, resulting in relatively stable time and cost proportions. Scrum demonstrates a flexible but measurable investment in quality, where verification efforts are embedded within each sprint and represented as an incremental increase in effort and budget. In contrast, evolutionary prototyping lacks clearly defined boundaries for quality-related activities, leading to greater variability and uncertainty in both time and cost estimations. This comparison reinforces the importance of selecting a development methodology that aligns not only with technical requirements but also with an organization's capacity to plan, control, and sustain quality assurance efforts within defined time and budget constraints.

Overall, this study demonstrates that methodologies with well-defined and continuous quality assurance practices provide better support for controlling the Time of Quality and the Cost of Quality. Therefore, selecting an appropriate software development methodology is a strategic decision that should consider not only development speed or flexibility, but also the organization's capacity to plan, measure, and manage quality-related activities within defined time and budget constraints.

REFERENCES

- [1] P. Lledó and G. Rivarola, *Gestión de proyectos: cómo dirigir proyectos exitosos, coordinar los recursos humanos y gestionar el riesgo*. Buenos Aires, Argentina: Prentice Hall, 2007..
- [2] Project Management Institute, *A Guide to the Project Management Body of Knowledge (PMBOK Guide)*. Campus Boulevard, Pennsylvania, USA: Project Management Institute, Inc., 2013.
- [3] "IBM - Costa Rica," IBM, 2017. [Online]. Available: <https://www.ibm.com/cr-es/>
- [4] J. Wiley and S. Wiley, *Software Testing and Analysis*, 1st ed., 2008.
- [5] P. Mahmood and M. Beg, "Measuring Cost of Quality (CoQ) on SDLC projects is indispensable for effective Software Quality Assurance," *International Journal of Soft Computing and Software Engineering*, vol. 2, no. 9, pp. 1–15, 2012.
- [6] "ASQ - A Global Leader in Quality Improvement & Standards," ASQ, 2017. [Online]. Available: <https://asq.org/>
- [7] R. Pressman and B. Maxim, *Software Engineering*, 1st ed. New York, NY, USA: McGraw-Hill, 2015.
- [8] K. Schwaber and J. Sutherland, *Software in 30 Days*, 1st ed. New York, NY, USA: John Wiley & Sons, 2012.
- [9] "World Quality Report 2016-17," Capgemini Worldwide, 2017. [Online]. Available: https://www.theseus.fi/bitstream/handle/10024/135804/Williams_Paivi.pdf?sequence=1&isAllowed=y
- [10] K. Schwaber and J. Sutherland, *The Scrum Guide*, 2020.
- [11] I. Sommerville, *Software Engineering*, 10th ed. Pearson, 2016.

TRANSFORMACIÓN DIGITAL E INTELIGENCIA ARTIFICIAL: POR QUÉ LA ADOPCIÓN TECNOLÓGICA NO GARANTIZA RESULTADOS.

PLANIFICACIÓN, DATOS Y MADUREZ ORGANIZACIONAL COMO FACTORES CRÍTICOS

Ing. J. Alonso Solano Segura LTI.

*Autor. Docente Universitario, San José, Costa Rica
asolanos@uned.ac.cr*

RESUMEN

La transformación digital y la inteligencia artificial se han promovido como soluciones capaces de generar mejoras automáticas en eficiencia e innovación; sin embargo, la evidencia demuestra que la adopción tecnológica por sí sola no garantiza resultados sostenibles. La digitalización de procesos mal definidos y el uso de datos de baja calidad suelen amplificar problemas organizacionales preexistentes, lo cual conduce al fracaso de iniciativas digitales. Desde una perspectiva crítica, se discute que la transformación digital es un proceso humano y estratégico que requiere planificación, madurez institucional y una gestión rigurosa de los datos. Además, se analiza la dependencia de la inteligencia artificial respecto a la calidad y gobernanza de la información, así como la necesidad de validación en su aplicación. La discusión integra literatura académica reciente y evidencia empírica para demostrar que el valor de la tecnología emerge únicamente cuando negocio, personas y capacidades tecnológicas actúan de forma integrada y coherente.

Palabras clave: Transformación digital; Inteligencia artificial; Planificación estratégica; Gobernanza de datos; Madurez organizacional

ABSTRACT

Digital transformation and artificial intelligence have been widely promoted as solutions capable of generating automatic improvements in efficiency and innovation; however, evidence shows that technological adoption alone does not guarantee sustainable results. The digitalization of poorly defined processes and the use of low-quality data often amplify pre-existing organizational problems, leading to the failure of digital initiatives. From a critical perspective, digital transformation is understood as a human and strategic process that requires planning, organizational maturity, and rigorous data management. Furthermore, the structural dependence of artificial intelligence on data quality and governance is examined, as well as the need for continuous human validation in its application. The discussion integrates recent academic literature and empirical evidence to demonstrate that technological value emerges only when business, human, and technological capabilities are aligned and integrated in a coherent manner.

Keywords: Digital transformation; artificial intelligence; strategic planning; data governance; organizational maturity.

INTRODUCCIÓN

La transformación digital (TD) se ha consolidado como una prioridad estratégica en organizaciones de todos los sectores, impulsada por la promesa de mejoras sustanciales en eficiencia operativa, innovación y competitividad. No obstante, en la práctica, muchas organizaciones caen en la trampa de “digitalizar el caos”, es decir, aplicar plataformas tecnológicas emergentes o soluciones de inteligencia artificial (IA) sobre procesos mal definidos, estructuras organizacionales inmaduras y datos de baja calidad. La evidencia empírica demuestra que esta aproximación rara vez produce los resultados esperados: se estima que cerca del 70 % de las iniciativas de TD no logran cumplir sus objetivos

estratégicos [1]. Las elevadas inversiones en tecnología, lejos de traducirse automáticamente en mejoras sostenibles, suelen derivar en resultados decepcionantes cuando no van acompañadas de cambios organizacionales profundos.

En este contexto, se plantea el problema central que motiva este ensayo: la tendencia recurrente de muchas organizaciones a implementar tecnología sin bases sólidas, con la expectativa de obtener resultados casi milagrosos. Esta visión simplista, frecuente en algunos niveles directivos, asume que la sola adquisición de un nuevo software, la proclamación de una TD o la incorporación puntual de IA en un área específica bastan para mejorar el desempeño organizacional. A ello se suma la

influencia de ciertos discursos promovidos por consultoras y medios especializados, que en ocasiones presentan estas iniciativas como soluciones prácticamente automáticas o “mágicas”. Sin embargo, la experiencia y la literatura académica coinciden en que la tecnología por sí sola no genera valor: tanto la TD como la IA dependen críticamente de la planificación estratégica, la madurez organizacional y la calidad y gobernanza de los datos para producir impactos reales y sostenibles.

En este artículo, se analiza la premisa central de que la TD y la fracasan cuando se conciben como soluciones tecnológicas aisladas y no como procesos integrales que involucran estrategia, organización y datos de calidad. Para sustentar esta tesis, primero se conceptualiza la TD como un fenómeno eminentemente organizacional (no meramente tecnológico). En segundo lugar, se explora la naturaleza de la IA y su dependencia estructural en los datos. En tercer lugar, se analiza la importancia de la planificación estratégica y la madurez organizacional como cimientos de cualquier iniciativa digital exitosa. En cuarto lugar, se revisan evidencias de fracasos frecuentes en enfoques tecnologicistas, de modo que se identifican sus causas. En la discusión se integran hallazgos de la literatura reciente para reforzar la premisa, y finalmente se presentan conclusiones y recomendaciones para organizaciones públicas y privadas.

TD COMO FENÓMENO ORGANIZACIONAL

La TD es un proceso amplio de integración de tecnologías digitales en todos los aspectos del negocio (operaciones, productos, servicios y modelo de negocio), con el fin de mejorar el desempeño organizacional y generar valor. No se trata solo de tecnología, sino de repensar cómo la empresa funciona en la era digital, implicando cambios en procesos, estructuras, cultura y competencias del personal. Como señalan Vărzaru y Bocean [2], la mera implementación de nuevas tecnologías no garantiza el éxito; las empresas deben replantear fundamentalmente sus modelos de negocio para aprovechar plenamente el potencial digital [2]. En otras palabras, la TD requiere un giro estratégico en la forma en que la organización genera valor, se relaciona con sus clientes y asegura su ventaja competitiva, más allá de la simple adopción de herramientas tecnológicas. Este cambio conlleva innovaciones organizativas que redefinen prácticas empresariales tradicionales y abren oportunidades de crecimiento inéditas.

Es útil definir la madurez digital u organizacional como el grado en que una organización ha desarrollado las capacidades internas necesarias para aprovechar eficazmente las tecnologías digitales. Una organización con alta madurez digital cuenta con procesos bien gestionados, liderazgo comprometido, estrategia clara, talento capacitado, datos gobernados y cultura de mejora continua. Por el contrario, en empresas de baja madurez predominan procesos manuales o fragmentados, datos aislados y pobre coordinación, lo cual dificulta cualquier iniciativa tecnológica. Diversos estudios han propuesto modelos para evaluar la madurez digital. Por ejemplo, ciertos marcos evalúan dimensiones como estrategia, calidad de datos, infraestructura tecnológica, gobierno de datos y cultura organizacional, lo cual evidencia que no solo importan las capacidades técnicas, sino también la cultura y los procesos de la organización para que la transformación tenga éxito [3]. En suma, la TD es esencialmente un proceso de cambio organizacional habilitado por tecnología: implica reestructurar procesos internos, adoptar métodos de trabajo ágiles y colaborativos, y desarrollar las competencias digitales de la fuerza laboral [2]. Las empresas que logran esta integración holística de tecnología con estrategia y organización estarán mejor posicionadas para adaptarse al entorno dinámico y liderar en la economía digital.

La literatura reciente refuerza esta visión integradora. Neacșu y Popescu [4], en un análisis bibliométrico, observaron una creciente convergencia entre las tecnologías digitales (como IA) y las estrategias de negocio durante 2021-2025 [4]. Esto subraya que las organizaciones más avanzadas no abordan la tecnología de forma aislada, sino alineada estrechamente con sus objetivos estratégicos. En consecuencia, la TD efectiva exige liderazgo estratégico: no es un proyecto de TI, sino un programa de cambio organizacional guiado desde la dirección y apoyado por toda la estructura empresarial. Esta perspectiva concuerda con hallazgos previos como el de Kane et al., quienes concluyeron que “la estrategia, no la tecnología, impulsa la TD” [5]. En esencia, el valor emerge cuando la tecnología se implementa con un propósito claro, coherente con la estrategia corporativa y apoyado por una cultura organizacional preparada para el cambio.

IA Y DEPENDENCIA DE LOS DATOS

La IA se refiere a sistemas informáticos capaces de realizar tareas que típicamente requieren inteligencia humana, como el aprendizaje, la predicción, el reconocimiento de patrones o la toma de decisiones, mediante algoritmos entrenados con grandes volúmenes de datos. En el contexto empresarial actual, la IA se vislumbra como un catalizador de la TD, habilitando la automatización avanzada, el análisis masivo de información y la personalización de productos/servicios a escala. No obstante, la IA depende críticamente de los datos: la calidad, cantidad y relevancia de los datos que alimentan los algoritmos determinan en gran medida el rendimiento de las soluciones de IA.

Un principio fundamental es “garbage in, garbage out”: datos deficientes producen resultados deficientes. La investigación sobre data-centric AI (IA centrada en datos) ha surgido precisamente debido a que históricamente se ha puesto más énfasis en mejorar los modelos que en asegurar buenos datos. Jarrahi et al. destacan que se tendió a privilegiar el modelo al costo de la calidad de los datos, y que los problemas en los datos han minado el desempeño de muchos sistemas de IA en aplicaciones reales [6]. De hecho, la falta de atención a la calidad de los datos ha derivado en resultados decepcionantes en proyectos de IA, especialmente en dominios de alto riesgo como la salud, lo que motivó un cambio de paradigma hacia enfoques centrados en mejorar sistemáticamente los datos tanto como los algoritmos [7]. En otras palabras, incluso el modelo de IA más avanzado no logrará “magia”, si la base de datos sobre la que opera es pobre o caótica.

La gobernanza de datos se vuelve entonces un pilar imprescindible para iniciativas de IA exitosas. Esto abarca asegurar la calidad (exactitud, consistencia, integridad), la cantidad suficiente de datos relevantes, así como el acceso y silo adecuado de la información en la organización. Las empresas líderes invierten en limpiezas de datos, integración de fuentes diversas y generación de datos de alto valor, entendiendo que sus algoritmos solo serán tan buenos como los datos que los entrenan. Por ejemplo, en el sector financiero y de marketing digital, se ha evidenciado que una mejora modesta en la calidad de datos puede aumentar significativamente la precisión de modelos predictivos [8].

Este principio también se reconoce a nivel macro: los datos son un activo estratégico nacional en la era de la IA. El informe Technology and Innovation Report 2025 de UNCTAD [7] advierte que, para aprovechar plenamente la IA, los países deben invertir en infraestructura digital, desarrollar talento especializado y, crucialmente, consolidar bases de datos de calidad. No es casualidad que menos de un tercio de los países en desarrollo cuenten actualmente con una estrategia nacional de IA [7]; muchos carecen de las bases de datos robustas y accesibles necesarias para entrenar sistemas de IA avanzados, lo cual los pone en desventaja. De este modo, puede afirmarse que la IA no es una solución plug-and-play: su implementación eficaz requiere construir cimientos sólidos de datos. Sin datos confiables y bien gestionados, la IA está condenada a producir resultados erráticos o sesgados. Por el contrario, con datos abundantes, limpios y pertinentes, la IA puede revelar todo su potencial para generar conocimientos accionables y apoyar la toma de decisiones informada.

PLANIFICACIÓN ESTRATÉGICA Y MADUREZ ORGANIZACIONAL

Antes de embarcarse en proyectos de TD o de IA, las organizaciones deben alcanzar cierto nivel de madurez organizacional y realizar una planificación estratégica rigurosa. Esto implica, en primera instancia, llevar a cabo un diagnóstico honesto de la situación actual: cuáles son las fortalezas y debilidades en procesos, cultura, estructura y recursos de datos de la empresa. Numerosas iniciativas digitales fracasan por haberse lanzado sin un entendimiento claro del punto de partida y sin objetivos realistas. Por tanto, una planificación estratégica efectiva en TD comprende definir una visión alineada con la estrategia de negocio, establecer metas concretas (p. ej., mejorar la eficiencia en X %, reducir tiempos de ciclo, aumentar la satisfacción del cliente, etc.) y elaborar una hoja de ruta que priorice iniciativas y recursos.

La madurez organizacional juega aquí un rol central. Una organización madura en términos digitales cuenta con procesos estandarizados y optimizados, una estructura flexible, liderazgo transformacional, personal capacitado y una cultura proclive a la innovación y al uso de datos en la toma de decisiones. En tal contexto, las iniciativas tecnológicas tienen un terreno fértil donde arraigar. Por el contrario, si una empresa de baja madurez intenta implementar IA o plataformas digitales sin antes fortalecer sus fundamentos, corre un alto riesgo de fracasar. Como lo resume coloquialmente el artículo base: “la tecnología no

sustituye el orden, solo lo amplifica”; digitalizar procesos caóticos solo genera caos digital más rápidamente. Por ello, primero se deben “hacer las cosas bien, luego hacerlas digital”. En la práctica, esto significa que la organización debe invertir en gestión por procesos, calidad y cambio cultural antes o simultáneamente a la inversión tecnológica.

Un componente crítico de la planificación es el mapeo y optimización de procesos. Aplicando disciplinas como Business Process Management (BPM), las organizaciones pueden documentar sus flujos de trabajo, identificar ineficiencias y eliminar cuellos de botella. Un proceso bien definido (con entradas, actividades, responsables, salidas y métricas) es candidato para automatizar por tecnología con alta probabilidad de éxito. En cambio, si se digitaliza un proceso mal definido o innecesariamente complejo, la tecnología solo replicará esos problemas. Por eso, frameworks de calidad como ISO 9001 o metodologías Lean/Seis Sigma enfatizan “estabilizar y mejorar el proceso antes de automatizarlo”. Estudios recientes confirman que la TD efectiva suele ir precedida de iniciativas de mejora de procesos y desarrollo organizacional [2], [3]. La planificación estratégica de la transformación debería incluir quick wins (digitalizaciones de procesos ya optimizados para generar valor rápido), así como proyectos de mediano y largo plazo que estén condicionados a elevar previamente la madurez en ciertas áreas.

Otro factor es la gestión del cambio organizacional, introducir nuevas tecnologías sin gestionar adecuadamente el factor humano capacitaciones, comunicación, gestión de expectativas, incentivos suele provocar resistencia interna y subutilización de las herramientas. La alta dirección debe articular una visión convincente de la transformación, explicar el porqué del cambio y fomentar una cultura de adopción tecnológica. Neacșu y Popescu [4] sugieren que aprovechar plenamente la IA en los negocios implica no solo desplegar algoritmos, sino preparar a la organización para cambiar y crecer en el mundo digital [4]. Esto incluye desarrollar nuevas habilidades en la plantilla (competencias digitales, analíticas y de manejo de datos), redefinir roles y procesos de trabajo, y establecer gobernanza para supervisar la transformación (por ejemplo, oficinas de proyecto, comités de datos, marcos éticos para IA, etc.).

Cabe destacar que el apoyo institucional y el entorno también cuentan. A nivel país, como indica UNCTAD [7], muchos gobiernos carecen aún de políticas integrales para IA y TD, lo que puede dificultar a las empresas locales obtener talento o infraestructura adecuada [7]. La planificación

estratégica organizacional idealmente debe alinearse con iniciativas públicas, con lo cual se aprovechan incentivos, marcos regulatorios y colaboraciones academia-industria que puedan existir.

A modo de síntesis, se puede decir que la planificación estratégica rigurosa y la construcción de madurez organizacional son prerrequisitos para cosechar valor de la TD y la IA. La tecnología debe insertarse en una organización preparada: con objetivos claros, procesos ordenados, cultura adaptativa y datos confiables. Sin estos pilares, las inversiones digitales estarán construidas sobre cimientos débiles. Como advierte un estudio de Vărzaru [2], los hallazgos subrayan la necesidad de que las empresas integren estratégicamente las tecnologías digitales en todos los ámbitos del negocio para sostener su competitividad a largo plazo [2]. En otras palabras, la transformación no es un proyecto puntual de TI, sino un proceso continuo de desarrollo organizacional guiado por la estrategia.

LIMITACIONES DE LA TRANSFORMACIÓN BASADA SOLO EN TECNOLOGÍA

La historia reciente de la TD está repleta de casos fallidos o resultados por debajo de lo esperado, lo que brinda valiosas lecciones sobre qué no hacer. Una constante en estos fracasos es haber adoptado un enfoque tecnocrático; es decir, creer que bastaba con comprar o implementar cierta tecnología para resolver problemas de negocio, sin abordar los aspectos humanos, estratégicos y de datos. Diversas encuestas y análisis (HBR, McKinsey, BCG) convergen en que alrededor de 70 % de los proyectos de transformación organizacional no logran sus objetivos [1], [5]. A continuación, se sintetizan causas recurrentes de fracaso identificadas en la literatura:

- Falta de visión y liderazgo: Iniciativas lanzadas sin un norte claro o sin el respaldo decidido del liderazgo suelen naufragar. Cuando los ejecutivos delegan completamente la transformación al departamento de TI, sin integrar el esfuerzo al plan estratégico global, la iniciativa pierde impulso. McKinsey destaca que frecuentemente “el CEO no establece una aspiración suficientemente alta ni comunica una narrativa convincente del cambio”, lo que lleva a falta de convicción en el equipo [1].
- Resistencia cultural y escasa gestión del cambio: Muchas transformaciones tecnológicas topan con resistencia del personal que teme por sus puestos o no entiende los beneficios del cambio. Si la organización carece de una cultura de

innovación o aprendizaje, los empleados ven las nuevas herramientas como una imposición. Estudios reportan que la falta de buy-in de los empleados y de mecanismos formales de gestión del cambio (capacitación, comunicación, incentivos) es un factor principal en el fracaso [1]. Por ejemplo, implementar una plataforma digital sin entrenar adecuadamente a los usuarios típicamente resulta en baja adopción y retorno nulo de la inversión.

- **Déficit de competencias y recursos:** Un error común es subestimar las capacidades necesarias para ejecutar la transformación. Organizaciones que emprenden proyectos de IA sin contar con científicos de datos calificados, o sin liberar a personal clave de sus responsabilidades diarias para dedicarse al proyecto, suelen estancarse. La TD requiere nuevos perfiles (analistas de datos, ingenieros de procesos, expertos en experiencia de usuario, etc.) y a menudo implica redistribuir recursos internos. No proveer estos recursos es receta para el incumplimiento de metas [1].
- **Enfoque puramente tecnológico (falta de alineación con el negocio):** Muchos fracasos provienen de adoptar tecnologías de moda per se, sin un caso de negocio sólido. Por ejemplo, implantar una solución de Big Data o de machine learning sin haber definido previamente qué problema de negocio se intenta resolver o qué indicadores mejorarán. Esto conduce a soluciones desconectadas de la realidad empresarial, que no generan valor tangible. Varios autores advierten del riesgo de un “síndrome de la solución en busca del problema”. La transformación debe partir de un diagnóstico de problemas u oportunidades y luego evaluar qué tecnología es adecuada, no al revés.
- **Datos de mala calidad o silos de información:** Como se discutió, si la organización intenta volcarse a la IA o a la analítica sin antes haber asegurado la integridad y disponibilidad de sus datos, probablemente obtendrá resultados pobres. Existen casos documentados de proyectos de IA que fallaron por datos sesgados, incompletos o no representativos, lo cual llevó a conclusiones erróneas y pérdida de confianza de los directivos en la herramienta. Sin un buen gobierno de datos, la magia prometida por la IA se diluye [7].
- **Dependencia excesiva de proveedores externos:** Algunas organizaciones, al carecer de capacidades internas, entregan por completo su transformación a consultoras o vendedores tecnológicos. Si bien es válido buscar apoyo externo, la falta de involucramiento interno crea dependencia y soluciones poco ajustadas a la cultura de la empresa. Además, el conocimiento queda fuera y la organización no desarrolla ownership sobre sus nuevas herramientas. Esto puede conducir al fracaso una vez que el proveedor se retira o cuando surgen necesidades de adaptación que la empresa no sabe cubrir.

Desde la perspectiva de política pública, también hay evidencias de enfoques tecnologicistas fallidos. Por ejemplo, gobiernos que implementaron plataformas de smart city o de gobierno digital sin una adecuada preparación institucional han visto sistemas infrautilizados o abandonados. La ONU (UNCTAD) [7] advierte que la IA puede exacerbar brechas de desarrollo si no se acompaña de políticas inclusivas y cooperación internacional, pues sus beneficios tienden a concentrarse en unos pocos actores líderes [7]. Es decir, concebir la IA como panacea tecnológica sin marcos de gobernanza amplios puede conducir al fracaso en lograr desarrollo equitativo.

En suma, la evidencia empírica y los análisis de casos dejan claro que un enfoque que vea la TD/IA solo como implantación de tecnología está condenado al tropiezo. Los fracasos reiterados refuerzan la premisa central de este ensayo: sin planificación estratégica, sin madurez organizacional y sin datos de calidad, la tecnología por sí misma no genera valor sostenible. Lecciones aprendidas de estos fracasos son incorporadas por organizaciones exitosas para “revertir la receta”: enfocarse primero en estrategia, procesos, personas y datos, y solo entonces habilitar con tecnología lo cual veremos a continuación en la discusión de buenas prácticas y evidencia positiva.

DISCUSIÓN

Tras analizar los fundamentos teóricos y la evidencia de campo, queda robustamente sustentada la idea de que la TD y la IA deben abordarse como procesos integrales de cambio organizacional, y no como simples compras de tecnología. Los hallazgos de la literatura científica respaldan esta visión de varias maneras complementarias.

En primer lugar, se enfatiza el rol de la estrategia y el liderazgo. Värzaru et al. [2] concluyen que las empresas deben integrar estratégicamente las tecnologías digitales en todas las facetas de su negocio para sostener ventajas competitivas a largo plazo [2]. Esto implica que la TD debe estar guiada por la estrategia corporativa: cualquier iniciativa tecnológica ha de responder a objetivos de negocio claros (crecimiento, eficiencia, satisfacción del cliente, innovación, etc.). Cuando esta alineación ocurre, la tecnología se convierte en un catalizador poderoso de la estrategia; de lo contrario, se vuelve un gasto sin dirección. El liderazgo ejecutivo tiene la responsabilidad de articular dicha alineación, priorizando proyectos digitales que generen valor y patrocinando el cambio cultural interno necesario.

En segundo lugar, la discusión resalta la importancia de la cultura organizacional y el capital humano. Los fracasos analizados mostraron que, sin la adhesión de las personas, las mejores herramientas serán infrautilizadas. Por ende, una transformación exitosa conlleva transformar a las personas sus habilidades, mentalidades y formas de trabajar en paralelo a la tecnología. Esto es consistente con la noción de organización ambidiestra que aprende y se adapta continuamente. La IA, en particular, debe implementarse de forma que potencie a los trabajadores en lugar de reemplazarlos. Los empleados necesitan comprender que la IA puede ser un segundo cerebro que les ayuda a tomar mejores decisiones y, para ello, se requieren programas de formación y una comunicación clara de beneficios. Un enfoque centrado en las personas también significa establecer gobernanza ética de la IA (para abordar sesgos, privacidad, impacto en puestos de trabajo), lo cual forma parte de la madurez organizacional en la era digital.

En tercer lugar, queda reafirmada la dependencia estructural de la IA respecto a los datos. La evidencia recopilada (sección II) muestra que la calidad de datos es frecuentemente el factor x que determina el éxito o fracaso de un proyecto de IA. Esto tiene implicaciones prácticas: las organizaciones deberían invertir tanto en mejorar sus datos (depuración, integración, enriquecimiento) como invierten en desarrollar o adquirir algoritmos. Incluso, se argumenta que en muchos casos es más rentable mejorar los datos existentes que buscar algoritmos más sofisticados [7]. Un ejemplo reciente es el auge de herramientas de AutoML y data-centric AI que ayudan a ingenieros a detectar y corregir problemas de datos de entrenamiento para elevar la precisión de modelos sin tocar el código. Además, gestionar bien los datos internamente (romper silos departamentales, unificar fuentes en lagos de datos, establecer responsables de calidad de datos) crea una ventaja informacional sostenible que trasciende cualquier tecnología específica.

En cuarto lugar, la discusión integra evidencia positiva de que cuando se sientan las bases correctas, la tecnología sí genera valor real. Por ejemplo, un estudio empírico reciente de Cui [8] en más de 6300 observaciones de empresas industriales chinas halló que la TD impulsada por IA aumenta significativamente el rendimiento de las empresas, pero dicho efecto se materializa plenamente solo cuando está mediado por la innovación organizacional y potenciado por la colaboración humano-IA en la empresa [8]. En otras palabras, la IA por sí misma podía mejorar la productividad,

pero las mayores ganancias ocurrían en empresas que simultáneamente innovaban en sus procesos (green digital innovation) y en las que los empleados trabajaban coordinadamente con las soluciones de IA (human-AI collaboration). Este hallazgo cuantitativo refuerza la tesis: la tecnología rinde frutos cuando se inserta en una organización innovadora y preparada para aprovecharla. Caso contrario, su impacto es limitado. Asimismo, Vărzaru et al. [2] han mostrado que la adopción de IA, Big Data, IoT y otras tecnologías digitales explica en parte el incremento de ingresos por innovación en empresas europeas, siempre y cuando dichas tecnologías estén integradas estratégicamente en la cadena de valor y acompañadas de cambios en el modelo de negocio [2]. Nuevamente, estrategia, organización y tecnología actuando en conjunto marcan la diferencia.

En quinto lugar, desde la perspectiva de políticas y desarrollo, se discute la necesidad de enfoques holísticos e inclusivos para que la tecnología tenga impacto social positivo. El informe de UNCTAD [7] recalca que la IA, vista aisladamente, no es inherentemente inclusiva ni garante de desarrollo [7]. Para que países en desarrollo se beneficien, deben simultáneamente construir infraestructura digital, formar talento local y diseñar marcos de gobernanza apropiados. Esta es la contracara a nivel macro de lo que a nivel micro ocurre en las organizaciones: sin capacitación (skills), sin infraestructura y sin gobernanza, la tecnología no obra milagros. UNCTAD [7] incluso llama a “cambiar el foco de la tecnología a las personas” y a una cooperación internacional para co-crear un marco global de IA centrado en la equidad [7]. Esto refleja una comprensión madura de que los beneficios de la TD deben orquestarse, no se generan automáticamente con la difusión tecnológica.

Si bien la literatura crítica enfatiza que la simple adopción tecnológica no garantiza resultados sostenibles, evidencia empírica reciente muestra que la TD puede impactar positivamente el desempeño empresarial cuando se combinan capacidades tecnológicas con orientación digital y cultura organizacional. Por ejemplo, estudios han encontrado que las capacidades de tecnologías de información están directamente correlacionadas con mejoras en desempeño organizacional en contextos empresariales reales, sugiriendo que la tecnología sí puede ser un motor de valor competitivo cuando se gestiona adecuadamente [10], [11], [12]. Estos y otros diversos estudios se pueden presentar como una antítesis al menos parcialmente ya que reportan efectos positivos de la TD en el desempeño empresarial; sin embargo, estos beneficios se observan principalmente en

organizaciones que ya cuentan con capacidades tecnológicas, orientación digital y estructuras organizacionales adecuadas [10], [11]. Lo anterior refuerza la idea de que la tecnología no actúa de forma autónoma, sino como habilitador dentro de un sistema organizacional maduro.

Tomando en conjunto todos estos puntos, se puede afirmar que existe un consenso amplio en la academia y organismos especializados sobre la premisa fundamental de la cual se viene investigando. Las organizaciones deben abordar la TD y la IA con una visión sistemática: tecnología + estrategia + organización + datos. Cualquier enfoque parcial o simplista, por ejemplo, centrar todos los esfuerzos en la herramienta tecnológica sin invertir en el factor humano, o viceversa, estará incompleto y probablemente falle. La discusión teórica y empírica aquí presentada proporciona un sustento sólido para dicha afirmación, a la vez que orienta sobre cómo proceder para maximizar las probabilidades de éxito en iniciativas digitales.

CONCLUSIONES

En la era contemporánea, saturada de promesas tecnológicas, se reafirma que la tecnología por sí sola no hace magia. La TD y la IA pueden, sin duda, generar enormes saltos en productividad, innovación y valor, pero solamente lo logran cuando se apoyan en cimientos firmes. Esos cimientos son: una planificación estratégica clara, una organización madura y adaptable, y datos de alta calidad bien gobernados. Sin estos elementos, los proyectos digitales se convierten en castillos de naipes. Con ellos, en cambio, la tecnología actúa como multiplicador y acelerador del cambio positivo.

A modo de cierre, y basado en los hallazgos de esta investigación, se proponen las siguientes recomendaciones clave para organizaciones (y, por analogía, gobiernos) que emprenden la senda de la TD e IA:

Definir una estrategia digital integral: Definir una visión y objetivos concretos de negocio que la tecnología ayudará a alcanzar. Alinear cada iniciativa tecnológica con esa estrategia, evitando proyectos aislados sin justificación en valor.

- **Diagnosticar y elevar la madurez organizacional:** Evaluar el estado actual de procesos, cultura, talento y datos. Fortalecer las áreas débiles mediante programas de mejora de procesos (BPM, Lean), gestión de calidad, capacitación digital del personal, documentación de procesos y cambio cultural hacia la innovación y la colaboración.

- **Establecer una gobernanza de datos y tecnología:** Implantar políticas y estructuras que aseguren la calidad, integridad, seguridad y accesibilidad de los datos, así como mecanismos de supervisión de los proyectos digitales (oficinas de transformación, comités de datos e IA y marcos éticos). La IA, la inteligencia de negocios y los algoritmos analíticos dependen críticamente de la calidad de los datos que los alimentan; cuando estos sistemas se basan en datos inconsistentes, no normalizados o de baja calidad, los resultados obtenidos son inevitablemente deficientes, comprometiendo la toma de decisiones y el valor esperado de las iniciativas digitales. En este sentido, los datos deben ser gestionados como un activo estratégico transversal de la organización.
- **Fomentar la participación y el desarrollo del talento:** Invertir en la formación continua del personal no solo en nuevas tecnologías, sino también en competencias transversales como gestión de procesos, calidad, gobierno y manejo de datos. La capacitación integral, acompañada de una comunicación clara sobre los objetivos de la transformación, facilita el compromiso organizacional (buy-in) y permite empoderar equipos multifuncionales capaces de liderar iniciativas digitales de manera efectiva, integrando conocimiento del negocio con capacidades de TI e IA.
- **Comenzar pilotos enfocados y escalables:** Iniciar con proyectos acotados que aborden pequeñas victorias (quick wins) alineados a la estrategia, para generar aprendizajes y demostrar valor. Usar estos casos de éxito iniciales para refinar la hoja de ruta e ir escalando gradualmente las soluciones digitales al resto de la organización.
- **Medir y aprender continuamente:** Definir indicadores clave (KPI) para cada iniciativa digital (por ejemplo, reducción de tiempo de proceso, aumento de satisfacción del cliente, incremento en ventas por canales digitales, etc.). Dar seguimiento constante a estos indicadores y estar dispuesto a iterar o recalibrar la estrategia en función de lo aprendido.
- **Reconocer el carácter humano de la TD:** La TD es fundamentalmente un proceso humano y organizacional. A pesar de los avances de la IA, su efectividad continúa dependiendo de la validación, supervisión y criterio humano. La tecnología no genera valor de forma autónoma; solo mediante un trabajo integral entre el negocio, el recurso humano y las capacidades tecnológicas es posible obtener resultados sostenibles y alineados con los objetivos organizacionales.

En conclusión, la TD y la IA deben concebirse como un viaje estratégico y organizacional, no como un destino tecnológico. Aquellas organizaciones que entiendan esto y actúen en consecuencia planificando con visión, de modo que cimentan ordenadamente sus procesos y datos, e involucrando a su gente en el cambio estarán en posición de convertir la tecnología en verdaderos resultados mágicos en términos de valor público o empresarial. Por el contrario, quienes busquen atajos tecnocráticos, esperando que la mera adquisición de IA o sistemas digitales resuelva problemas estructurales, continuarán engrosando la estadística de fracasos. La tecnología tiene un inmenso potencial transformador, pero es la combinación con liderazgo, estrategia, carácter humano y datos lo que finalmente determina el éxito o el caos.

REFERENCIAS

- [1] H. Robinson, "Why Do Most Transformations Fail? A Conversation with Harry Robinson," McKinsey & Company, 2019. [En línea]. Disponible en: <https://www.mckinsey.com/~media/McKinsey/Business%20Functions/Transformation/Our%20Insights/Why%20do%20most%20transformations%20fail%20A%20conversation%20with%20Harry%20Robinson/Why-do-most-transformations-fail-a-conversation-with-Harry-Robinson.pdf>
- [2] A. A. Vărzaru y C. G. Bocean, "Digital Transformation and Innovation: The Influence of Digital Technologies on Turnover from Innovation Activities and Types of Innovation," *Systems*, vol. 12, no. 9, art. 359, 2024. [En línea]. Disponible en: <https://www.mdpi.com/2079-8954/12/9/359>
- [3] A. Solano Segura, "Transformación digital e inteligencia artificial: la tecnología no hace magia," *Dos Tecnología y Negocios*, 2025. [En línea]. Disponible en: <https://www.dostecnologiaynegocios.com/2025/11/transformacion-digital-e-inteligencia.html>
- [4] N. A. Neacșu y A. Popescu, "Exploring the Impact of Artificial Intelligence on Digital Transformation of Businesses: A Bibliometric Analysis," 2025. [En línea]. Disponible en: https://www.researchgate.net/publication/394301486_Exploring_the_Impact_of_Artificial_Intelligence_on_Digital_Transformation_of_Businesses_A_Bibliometric_Analysis
- [5] G. C. Kane, D. Palmer, A. N. Phillips, D. Kiron y N. Buckley, "Strategy, Not Technology, Drives Digital Transformation," *MIT Sloan Management Review*, 2015. [En línea]. Disponible en: <https://sloanreview.mit.edu/projects/strategy-drives-digital-transformation/>
- [6] M. H. Jarrahi, A. Memariani y S. Guha, "The Principles of Data-Centric AI (DCAI)," *Communications of the ACM*, vol. 66, no. 1, pp. 40-48, 2023. [En línea]. Disponible en: https://www.researchgate.net/publication/365358551_The_Principles_of_Data-Centric_AI_DCAI
- [7] United Nations Conference on Trade and Development (UNCTAD), "AI's \$4.8 Trillion Future: UN Trade and Development Alerts on Divides, Urges Action," 2024. [En línea]. Disponible en: <https://unctad.org/press-material/ais-48-trillion-future-un-trade-and-development-alerts-divides-urges-action>
- [8] J. Cui, "AI-Driven Digital Transformation and Firm Performance in Chinese Industrial Enterprises: Mediating Role of Green Digital Innovation and Moderating Effects of Human-AI Collaboration," arXiv preprint, 2025. [En línea]. Disponible en: https://www.researchgate.net/publication/391878221_AI-Driven_Digital_Transformation_and_Firm_Performance_in_Chinese_Industrial_Enterprises_Mediating_Role_of_Green_Digital_Innovation_and_Moderating_Effects_of_Human-AI_Collaboration
- [9] B. Tabrizi, E. Lam, K. Girard y V. Irvin, "Digital Transformation Is Not About Technology," *Harvard Business Review*, mar. 2019. [En línea]. Disponible en: <https://hbr.org/2019/03/digital-transformation-is-not-about-technology>
- [10] V. Barba-Sánchez y C. Atienza-Sahuquillo, "Effects of digital transformation on firm performance," *International Journal of Information Management Data Insights*, vol. 4, no. 1, 2024. [En línea]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2405844024037563>
- [11] R. Masoud y S. Basahel, "The Effects of Digital Transformation on Firm Performance: The Role of Customer Experience and IT Innovation," *Digital*, vol. 3, no. 2, pp. 109-126, 2023. [En línea]. Disponible en: <https://www.mdpi.com/2673-6470/3/2/8>
- [12] E. Chrzanowska, M. Chrzanowski y P. Zawada, "AI-powered digital transformation: organizational implementation insights," *Journal of Management Sciences*, 2024. [En línea]. Disponible en: <https://www.jomswsge.com/pdf-197019-119946?filename=AI-powered-digital-transf.pdf>

ÉTICA, COMUNICACIÓN Y MARCA PROFESIONAL: PILARES CLAVE PARA EL LIDERAZGO RESPONSABLE EN TECNOLOGÍA

Ensayo

Maikol Solís

Centro de Investigación en Matemática Pura y Aplicada, UCR, San José, Costa Rica

Fundador STATO Pro, San José, Costa Rica

maikol.solis@statopro.com / maikol.solis@ucr.ac.cr

Irene Jara

Consultora en Imagen y Marca Profesional, Colegio de Consultores en Imagen Pública, Ciudad de México

Creadora del Método PRESENCIA®, Costa Rica

irene@irenejara.com

Palabras clave: Inteligencia artificial responsable; explicabilidad de modelos; gobernanza de datos; habilidades de poder; marca profesional; storytelling con datos.

RESUMEN

El éxito profesional en el ámbito de la tecnología y la ciencia de datos trasciende la mera experticia técnica, requiriendo la convergencia de la ética, la comunicación estratégica y el liderazgo responsable. En un futuro cercano, los roles como analistas de datos e ingenieros fintech crecen aceleradamente. Asimismo, la región latinoamericana enfrenta brechas críticas en gobernanza ética y en habilidades de poder (antes habilidades blandas). Es por esto por lo que estos roles deben tener presente cómo mejorar su perfil profesional para ser más atractivos en un mercado laboral más demandante.

INTRODUCCIÓN

El panorama global del mercado laboral en los próximos años se define por tres dinámicas simultáneas: (1) creación acelerada de empleos en roles técnicos de inteligencia artificial (IA) y análisis de datos, (2) transformación radical de cómo se toman decisiones empresariales basadas en modelos algorítmicos, y (3) creciente escrutinio regulatorio y social sobre la equidad, privacidad y transparencia de estos sistemas.

Según el Informe “Future of Jobs 2025” del Foro Económico Mundial [1], analistas de datos especializados en big data, ingenieros del sector fintech y especialistas en machine learning figuran entre los diez roles de más rápido crecimiento a nivel mundial, con proyecciones de creación de 11 millones de empleos netos atribuibles a IA y procesamiento de información entre 2025-2030. Sin embargo, esta misma fuente identifica que 63 % de empleadores globales considera que la brecha de habilidades es la barrera número uno para la transformación digital, y que 39 % de las competencias existentes de los trabajadores se volverá obsoleta en el mismo período.

Entre las habilidades técnicas con más crecimiento en los próximos 5 años están: Inteligencia artificial y big data (87 % de los empleadores), redes y ciberseguridad (70 %) y conocimiento tecnológico (68 %). En contraposición, las habilidades de poder (o blandas) están en una posición muy predominante. Por ejemplo, el pensamiento creativo (66 %), resiliencia, flexibilidad y agilidad (66 %), curiosidad y aprendizaje continuo (61 %) y liderazgo e influencia social (58 %). En todos los casos, se recalca que existe un desajuste entre lo que se enseña en universidades y certificaciones técnicas actualmente. En particular, se destaca que el 50 % de los trabajadores ya han completado capacitaciones (comparado con el 41 % en 2023), y que 59 de cada 100 de los trabajadores globales necesitarán entrenamiento para 2030. De estos, 29 podrían ser capacitados en su rol actual, 19 podrían ser preparados y reasignados a otros puestos dentro de la organización, pero 11 no recibirían esa formación, quedando sus perspectivas de empleo en riesgo.

En Latinoamérica, esta tensión es particularmente aguda. Los estudiantes lideran el conocimiento en IA generativa con un crecimiento de inscripciones 2,2 veces superior al promedio global (425 % vs. 195 %), según el “Global Skills Report 2025” [2]. En particular, la región gradúa casi un millón de profesionales en carreras técnicas anualmente. Sin embargo, persisten tres brechas críticas ampliamente documentadas:

Gobernanza ética: A pesar de esfuerzos normativos emergentes en Brasil, Chile y Argentina, la región carece de marcos regulatorios robustos para IA responsable. Iniciativas legislativas latinoamericanas mencionan principios éticos “en pasada,” sin medidas operacionales concretas para garantizar equidad algorítmica, privacidad de datos o auditorías de sesgo [3], [4]. El Índice Latinoamericano de Inteligencia Artificial 2025 [5] revela brechas significativas en gobernanza entre todos los países líderes como Brasil, Chile, Colombia y México, y los rezagados como Bolivia, Ecuador o Paraguay.

Comunicación estratégica: El 70 % de las organizaciones latinoamericanas reconocen que la falta de habilidades digitales y comunicativas en sus equipos técnicos constituye una barrera significativa para adopción de tecnologías disruptivas [6], [7]. Esta “brecha de comunicación” se agrava porque programas educativos de ciencia de datos globalmente enfatizan programación y machine learning, pero ofrecen mínima formación en comunicación, colaboración y consideraciones éticas. En [8] se confirma, mediante minería de texto, que los programas de maestría en ciencia de datos se enfocan principalmente en habilidades de investigación, procesamientos de datos, tecnologías de la información, entre otras. Solamente pocas universidades especializadas priorizan temas en ética o estadística.

Liderazgo situacional: En el estudio de Arlenis y Ronald [9], se mezcló una revisión literaria sistemática con entrevistas a líderes de la industria latinoamericana. Aproximadamente el 40 % de los entrevistados admiten inseguridad al aplicar herramientas digitales avanzadas directamente, delegando a equipos técnicos o consultores externos. Esto deja en evidencia un liderazgo situacional y capacidad de influencia sin autoridad formal. Este retraso en el desarrollo de habilidades de liderazgo representa un “cuello de botella” para modelos organizacionales ágiles, sostenibles e innovadores en la región. Es decir, aunque se tienen profesionales técnicamente competentes y contribuyen significativamente a la ejecución de tareas, sus déficits de comunicación e inteligencia emocional frecuentemente generan dinámicas de equipo pobres, liderazgo inefectivo y reducción de innovación organizacional [10].

La diferenciación profesional en ciencias de la información y ciencia de datos en los próximos años no dependerá únicamente de maestría técnica, sino de la integración consciente de la capacidad ética y gobernanza responsable de la IA; la comunicación efectiva usando técnicas de storytelling; y liderazgo y gestión de impacto organizacional.

I. GOBERNANZA ÉTICA

A nivel global, el ecosistema de gobernanza de IA se ha estructurado alrededor de tres marcos internacionales principales: (1) EU AI Act focalizada en clasificación de riesgo; (2) UNESCO Recommendations on AI Ethics, que enfatiza inclusividad, justicia y prevención de daño; y (3) NIST AI Risk Management Framework e ISO/IEC 42001 centrado en la medición y mitigación de riesgo algorítmico.

En Latinoamérica, este panorama regulatorio se ha acelerado significativamente entre 2024-2025, con desarrollo simultáneo de marcos nacionales y estrategias regionales en Perú, Brasil, Colombia, Chile, Argentina y México. Por su parte, Costa Rica lanzó su Estrategia Nacional de IA 2024-2027 (ENIA), de manera que se convirtió en el primer país centroamericano con estrategia nacional dedicada, liderada por el Ministerio de Ciencia, Innovación, Tecnología y Telecomunicaciones (MICITT). Costa Rica no cuenta aún con ley específica de IA; sin embargo, la ley No. 23097 “Protección de Datos Personales” está en revisión para alinear con el estándar EU GDPR. Además, se está participando activamente en los comités de IA de la OCDE para desarrollar una legislación más específica [11]. Este cambio de paradigma lleva a los profesionales en datos a enfrentar algunas áreas de riesgo críticas. Se mencionaron solo tres: Sesgo algorítmico, privacidad y explicabilidad. Pero ciertamente hay muchos otros que contemplar.

A. Sesgo algorítmico y equidad

Los modelos de machine learning son, en esencia, reflejos del mundo que les proporcionamos a través de los datos de entrenamiento. Si estos datos históricos contienen o reflejan sesgos sistémicos preexistentes (relacionados con variables sensibles como el género, la raza, la etnia, la orientación socioeconómica o la capacidad), el modelo no solo hereda esos sesgos, sino que a menudo los amplifica y automatiza [12].

La problemática del sesgo se acentúa en América Latina debido a sus marcadas desigualdades socioeconómicas y la prevalencia de la informalidad. Tres ejemplos de esta problemática son:

Acceso al crédito y la exclusión financiera. Muchos bancos usan puntajes de crédito para clasificar la capacidad de crédito a partir de historiales provenientes del sector formal. En muchos países de la región, el sector informal supera el 50 % de la fuerza laboral. Este tipo de sesgos establece una “correlación” espuria entre la informalidad. Esto puede poner a familias con capacidad de pago, pero excluidas de un crédito formal. Esto implica un riesgo de que estas familias tomen créditos fuera del sistema financiero o financiados por actividades ilícitas.

Contratación y recursos humanos: Los sistemas automatizados de screening pueden desechar currículos provenientes de instituciones educativas en áreas de bajo ingreso o aplicar filtros de lenguaje que excluyen dialectos regionales, perpetuando barreras de acceso al empleo.

Justicia predictiva: El uso incipiente de la IA en la predicción de riesgo de reincidencia penal puede basarse en datos históricos policiales que están sesgados contra ciertas comunidades o barrios, lo que lleva a sentencias o libertades condicionales más severas para individuos pertenecientes a esos grupos.

Para mitigar sesgos de manera activa se necesitan tres estrategias: (1) Auditorías multidimensionales que usen métricas de equidad específicas (como paridad demográfica e igualdad de oportunidades) antes y después de que el modelo entre en funcionamiento; (2) La eliminación de sesgos en tres etapas: pre-procesamiento (ajuste de pesos o muestreo), in-procesamiento (eliminación del sesgo con modelos adversarios donde modelos contrarios penalizan la predicción de atributos sensibles), y post-procesamiento (ajuste de umbrales por subgrupo); y (3) Equipos multidisciplinarios, entendiendo que la equidad varía según el contexto y necesita la colaboración de sociólogos, expertos legales y comunidades afectadas, lo cual es especialmente importante en Latinoamérica, donde la diversidad laboral y cultural es diferente a los contextos donde se crearon los supuestos iniciales. En estudios científicos sobre COVID-19 se demostró que los modelos adversos pueden mantener la precisión ajustando por equidad de género, raza, salarios, etc. [13]. Actualmente, existen herramientas que se encargan de estos procedimientos más directamente, como el IBM AI Fairness 360 [14].

B. Privacidad, protección de datos y el desafío de la ia generativa

La gestión masiva de datos personales es el cimiento de la IA, pero representa un riesgo significativo. El profesional de datos debe navegar un panorama regulatorio complejo y en constante evolución. Regulaciones pioneras como el GDPR (Europa) y la CCPA (California) han sentado las bases para los derechos del ciudadano sobre sus datos. En América Latina, países como Chile, Argentina, Colombia y Brasil (con la LGPD) han promulgado o están modernizando sus leyes de protección de datos, imponiendo obligaciones estrictas sobre el consentimiento, el propósito del uso y la seguridad.

Aun así, la llegada de los modelos de lenguaje grandes (LLM) y la IA generativa ha intensificado los riesgos de privacidad. En particular, se pueden dar los ataques de inferencia de membresía, donde modelos entrenados en vastos conjuntos de datos públicos y privados son vulnerables a que terceros puedan determinar si un dato específico (por ende, la información privada de un individuo) fue parte del conjunto de entrenamiento [15]. Otra vulnerabilidad persistente es el llamado “inversión del modelo”, donde se le da suficiente memoria a la IA generativa para “recordar” literalmente partes de sus datos de entrenamiento, permitiendo que prompts o instrucciones particulares reconstruyan información personal sensible [16].

Aunque no es sencillo, se pueden aplicar estrategias para coadyuvar en este tipo de salidas de información. En general, la estrategia debe enfocarse en: (1) la inyección controlada de ruido a los datos originales; (2) la capacidad de compartir solo los parámetros de modelo en lugar del modelo completo; (3) recolección solo de los datos estrictamente necesarios para su anonimización y (4) acuerdos claros y transparentes con terceros sobre el uso y eliminación de los datos personales [17].

C. Explicabilidad y transparencia

El problema de la caja negra se refiere a la incapacidad de los modelos complejos, como las redes neuronales profundas o los modelos de ensambles, para ofrecer una justificación clara y entendible de sus predicciones. En dominios de alto riesgo, la falta de explicabilidad no es solo una molestia técnica, sino un impedimento ético y legal [18]. Esto podría anular procesos en justicia, crear malos diagnósticos en medicina o crear injusticias injustificadas en el sector bancario. Sin explicabilidad, la confianza pública y regulatoria se socava, perpetuando decisiones injustas y no auditables.

Diversas técnicas estadísticas de interpretabilidad permiten abrir la caja negra algorítmica. Los SHapley Additive exPlanations (SHAP) values asignan importancia a cada variable basándose en teoría de juegos cooperativos, proporcionando explicaciones locales y globales, con garantías matemáticas de consistencia y precisión local. Los modelos agnósticos locales generan versiones simples, como por ejemplo, una regresión lineal alrededor de un caso específico para aproximar el comportamiento del modelo complejo en esa vecindad, permitiendo explicar por qué se tomó una decisión particular independientemente del algoritmo subyacente. Adicionalmente, los análisis de importancia de variables (para modelos basados en árboles) y los gráficos de dependencia parcial permiten visualizar la influencia marginal de variables en predicciones. Estas técnicas agnósticas del modelo son de suma importancia para cumplir con requisitos regulatorios de explicabilidad en aplicaciones de alto riesgo [19].

II. COMUNICACIÓN ESTRATÉGICA

Una reciente investigación basada en análisis de 786 descripciones de puestos en ciencia de datos revela que comunicación está presente en 100 % de las posiciones, constituyendo la habilidad más demandada, seguida por resolución de problemas (92 %) y trabajo en equipo (91 %). Las habilidades de poder o blandas comprenden la mayor parte de los resultados, lo cual confirma que estas habilidades mejoran significativamente el desempeño y efectividad en el lugar de trabajo. Específicamente, se tiene que las habilidades de comunicación efectiva permiten a profesionales articular hallazgos claramente a todas las partes interesadas del negocio (stakeholders), ya sean técnicos como no técnicos. El estudio demuestra que, sin importar el nivel o el rol, las habilidades de poder son muy necesarias para cada profesional. En particular, se debe dar énfasis a la comunicación escrita y oral. La razón es simple pero profunda: los modelos carecen de valor; los modelos sofisticados carecen de valor si los stakeholders no entienden, no confían o no actúan en base a los hallazgos.

Como un ejemplo de esta habilidad, supongamos que un científico de datos reporta lo siguiente en una presentación: “El modelo alcanzó F1-score de 0,87 en validación cruzada, con AUC-ROC de 0,91”. Aunque la información es técnicamente correcta, es estratégicamente inútil para los stakeholders de una empresa. En este caso, las métricas técnicas por sí solas no comunican valor de negocio a los

grupos de interés. La misma información se puede traducir a algo más útil como: “Para evitar más pérdidas debido al riesgo de abandono de clientes, nuestro modelo identifica correctamente al 87 % de esta población con 13 % de falsos positivos. Una vez implementado el modelo, este permitiría retener \$2,1M en ingresos anuales con costo de \$600K, obteniendo un ROI (retorno de inversión) de 250 %”.

Nótese que en el ejemplo anterior se aplicaron técnicas de storytelling efectivo con datos. Esta sigue estructura narrativa clásica [20]:

- Gancho: Situación que importa a la audiencia.
- Conflicto: Problema específico cuantificado que genera preguntas.
- Análisis: evidencia técnica comunicada a través del descubrimiento de puntos clave.
- Resolución: Plan accionable que elimina el conflicto.
- Impacto: Consecuencias monetarias de aplicar la resolución.

En la práctica, lo que se requiere es ser capaz de generar una o dos visualizaciones simples, explicar analogías intuitivas, presentar números redondos y contar, a través de secuencias lógicas, procesos complejos. El trabajo del profesional es ser capaz de conectar métricas técnicas con los resultados esperados por los stakeholders.

Otro aspecto críticamente subestimado es comunicar transparentemente lo que el modelo no puede hacer, y por qué. En contexto de IA responsable, esta comunicación no es opcional, es gestión de riesgo profesional [21]. Un caso típico es cuando se ajusta un modelo y una población en riesgo es subrepresentada (como en el caso de los trabajadores informales). La institución debería no usar el modelo para ese segmento en particular y hacer revisiones manuales. En caso de querer usarlo, se debería estar consciente de aceptar una gran cantidad de errores por el modelo. Esta transparencia sobre limitaciones, sesgos conocidos y condiciones extremas del modelo no necesariamente es algo malo. Es profesionalismo y manejo responsable del riesgo y ética profesional.

Como se mencionó anteriormente, los marcos regulatorios sobre la ética con IA están siendo un requerimiento en las empresas. La comunicación sobre los alcances, la incertidumbre de los modelos usados es una forma ética de transparencia. Esto permite a los stakeholders, evaluar el nivel de confianza en los modelos, identificar situaciones donde las decisiones son inciertas o fuera

de parámetros de entrenamiento, y preservar el juicio humano para casos más extremos mientras se permite el uso de la IA para decisiones rutinarias [22]. No hacer esto podría provocar pérdida de confianza y dinero a las organizaciones, como ya ha ocurrido en el pasado [23].

III. LIDERAZGO, MARCA PERSONAL, Y CULTURA ORGANIZACIONAL

Los profesionales en el campo de la tecnología y la información enfrentan hoy un entorno altamente competitivo, dinámico y globalizado. Este exige no solo dominio técnico, sino competencias interpersonales sofisticadas. En la literatura, se demuestra que las habilidades de poder (blandas) comprenden 11 de las 21 habilidades más demandadas en posiciones de datos y tecnología [24]. De este estudio, se resaltan la comunicación, resolución de problemas, trabajo en equipo, colaboración, entre otras. En otro estudio con trabajadores portugueses, se demostró que el desarrollo de habilidades de poder presenta correlación positiva significativa con progresión de carrera [25]. Este hallazgo establece que, mientras las habilidades técnicas pueden asegurar esa contratación inicial, las habilidades de poder (particularmente comunicación, liderazgo y marca profesional) son las que determinan trayectoria de carrera, velocidad de ascenso y acceso a posiciones estratégicas. La imagen y marca profesional se convierten así en activos intangibles críticos para el crecimiento profesional, la empleabilidad sostenida y el acceso a posiciones de mayor responsabilidad e influencia. Esta marca profesional se construye sobre cuatro fundamentos interdependientes que, gestionados coherentemente, fortalecen la percepción de autoridad y liderazgo en entornos profesionales altamente técnicos.

A. Imagen personal estratégica

La imagen personal trasciende apariencia estética superficial. En contexto profesional, se refiere a la capacidad de alinear presentación personal con rol organizacional, entorno y objetivos de carrera. En tecnología, imagen personal adecuada comunica orden mental, profesionalismo, atención al detalle y respeto por el entorno laboral. Estos elementos influyen directamente en la percepción de confiabilidad y preparación para asumir mayores responsabilidades. Un metaanálisis [26] confirma que individuos forman impresiones iniciales sobre confianza, competencia y profesionalismo en menos de un segundo a exposición a fotografías de personas desconocidas. Estos

juicios están basados en señales como postura corporal, contacto visual, vestimenta y peinado. En contextos profesionales, estas impresiones rápidas influyen significativamente en decisiones de contratación, asignación de proyectos y evaluaciones de credibilidad.

B. Comunicación verbal y no verbal

La comunicación es el vehículo principal de habilidades de poder en el sector tecnológico. La capacidad de explicar conceptos técnicos complejos, argumentar decisiones de diseño y persuadir a grupos de interés sobre el negocio es tan relevante como la capacidad de ejecutar implementaciones técnicas. Una comunicación verbal clara, estructurada y orientada a resultados permite que el conocimiento técnico sea comprendido y valorado por audiencias no técnicas. Esta condición ayuda a que profesionales técnicos asciendan a roles de liderazgo donde deben traducir estrategia tecnológica en valor de negocio.

En paralelo, la comunicación no verbal, como postura corporal, contacto visual, tono de voz, gestos y proxémica, refuerza credibilidad y seguridad personal. También puede socavar la imagen cuando es inconsistente con el mensaje verbal. Estudios en psicología organizacional indican que gestos no verbales de dominancia (postura erguida, contacto visual directo, voz firme) están positivamente asociados con confianza, mientras que indicadores de tensión o nerviosismo (voz temblorosa, evitación de mirada, postura encogida) están negativamente asociados con la confianza percibida [27].

Específicamente, señales no verbales asociadas con mayor confianza incluyen: (1) cantidad elevada de contacto visual, (2) orientación corporal directa hacia interlocutor, (3) inclinación hacia adelante, (4) habla fluida sin pausas prolongadas, (5) latencias de respuesta cortas y (6) turnos de conversación prolongados que demuestran dominio del tema.

C. Comportamiento profesional y habilidades de poder

El comportamiento profesional incorpora reglas, posturas y comportamientos que reflejan madurez emocional, liderazgo ético y profesionalismo constante. Incorpora capacidades de autoridad tales como inteligencia emocional, asertividad, capacidad de tomar decisiones bajo presión, negociación eficaz y gestión constructiva de disputas. En entornos tecnológicos de gran tensión donde ocurren lanzamientos de productos, incidentes de producción, negociaciones con clientes, entre

otras, estas competencias posibilitan conservar autoridad sin necesidad de subordinación jerárquica, establecer vínculos de confianza duraderos y liderar desde influencia en vez de desde autoridad posicionada.

La inteligencia emocional está positivamente correlacionada con los aspectos valorados, como cultura organizacional, desempeño y satisfacción. Está inversamente relacionado con el estrés. Es decir, que líderes emocionalmente inteligentes presentan significativamente una mayor participación y satisfacción de parte de sus equipos y experimentan menos estrés [28].

D. Marca profesional y reputación digital

La marca profesional que se deja en el ambiente de trabajo y digital es la coherencia entre lo que hace, lo que comunica y cómo es percibido por otros. En tecnología, esta marca se construye tanto en espacios presenciales (conferencias, reuniones ejecutivas, mentoría de juniors) como digitales: LinkedIn, GitHub, blogs técnicos, participación en comunidades de código abierto y ponencias en eventos de industria. Además, en economías digitales donde interacciones profesionales ocurren frecuentemente en entornos virtuales, imagen personal estratégica incluye también presencia en videollamadas (fondo profesional, iluminación adecuada y cámara activada), firmas de correo electrónico consistentes, y fotografías profesionales actualizadas en perfiles digitales. Incluso, se sabe que el 93 % de empleadores revisan perfiles de redes sociales para evaluar la empleabilidad de candidatos. Estas revisiones son complementadas con el CV para tomar la decisión final de contratación [29].

CONCLUSIÓN

El éxito laboral en tecnología moderna requiere la convergencia entre experticia técnica, comunicación eficaz y responsabilidad moral. Se necesitan expertos que transmitan de manera clara las restricciones de modelos, los prejuicios reconocidos y los riesgos de discriminación, en vez de esconderlos en lenguaje técnico y complicado. Esto construye una reputación de confiabilidad y liderazgo consciente. Hay que ser enfáticos: la comunicación clara sobre dudas, explicabilidad, sesgos, compromisos y limitaciones operativas en los modelos de IA no debilita la credibilidad técnica, sino que la refuerza. Esto permite a los usuarios tomar decisiones informadas y mantener la confianza en la organización a largo plazo. Se ha visto cómo la comunicación está en la mayoría de las descripciones de trabajos en tecnología. Esto, aunque los centros de formación técnica no las imparten. Estas habilidades son igual de importantes que el conocimiento técnico para avanzar en la carrera profesional. Cuando la experiencia

técnica se une a una comunicación responsable y ética, surgen líderes que pueden dirigir a las organizaciones hacia la innovación, equilibrando el rendimiento técnico con el impacto en las personas y la sociedad.

Finalmente, se debe hacer hincapié en que la imagen y la marca profesional ayudan a aumentar el poder y la influencia en el trabajo. Esto hace que los conocimientos técnicos sean vistos, valorados y reconocidos dentro de la organización. Los profesionales que gestionan estratégicamente estos elementos son percibidos como líderes naturales incluso antes de ocupar cargos formales. Como punto central, se encuentran los líderes con inteligencia emocional alta, que se relacionan con mejores ambientes de trabajo y bajos niveles de estrés. Invertir en la marca profesional no es un acto superficial, sino una estrategia de carrera que permite cerrar la brecha entre competencia técnica y reconocimiento organizacional. Esto facilita el acceso a nuevas oportunidades, influencia en decisiones estratégicas y construcción de reputación que perdura en un entorno profesional altamente competitivo y globalizado.

REFERENCIAS

- [1] World Economic Forum, "Future of Jobs Report 2025," World Economic Forum, ene. 2025. [En línea]. Disponible en: https://reports.weforum.org/docs/WEF_Future_of_Jobs_Report_2025.pdf
- [2] Coursera, "Global Skills Report 2025," 2025. [En línea]. Disponible en: https://assets.ctfassets.net/2pudprfttv6/3ELFKTA8GBPBuRkNrOzmpS/24fc7ec2372d0adb96965340069f705c/Global_Skills_Report_2025.pdf
- [3] "Regional cooperation crucial for AI safety and governance in Latin America," Brookings. [En línea]. Disponible en: <https://www.brookings.edu/articles/regional-cooperation-crucial-for-ai-safety-and-governance-in-latin-america/>
- [4] V. Muñoz, "Operation regulation: Strengthening Latin America's AI governance," European Council on Foreign Relations (ECFR). [En línea]. Disponible en: <https://ecfr.eu/article/operation-regulation-strengthening-latin-americas-ai-governance/>
- [5] CEPAL, "Índice Latinoamericano de Inteligencia Artificial (ILIA) 2025," dic. 2025. [En línea]. Disponible en: <https://hdl.handle.net/11362/82514>
- [6] OECD, "2023 OECD/IDB Digital Government Index of Latin America and the Caribbean: Results and key findings," OECD Public Governance Policy Papers, no. 64, nov. 2024. doi: 10.1787/10b82c83-en.
- [7] A. Cathles, C. Suaznabar y F. Vargas, "Radiografía de la transformación digital en las firmas de América Latina y el Caribe: resumen," Inter-American Development Bank, jun. 2022. doi: 10.18235/0004328.

- [8] M. J. Mol, B. Belfi y Z. Bakk, "Unravelling the skills of data scientists: A text mining analysis of Dutch university master programs in data science and artificial intelligence," *PLoS One*, vol. 19, no. 2, p. e0299327, feb. 2024. doi: 10.1371/journal.pone.0299327.
- [9] M. Arlenis Isabel y L. Ronald, "Digital Leadership in Latin America: Key Competencies for Managers," *Publis Editorial*, jul. 2025. doi: 10.5281/ZENODO.16753840.
- [10] F. S. Mohammed y F. Ozdamli, "A Systematic Literature Review of Soft Skills in Information Technology Education," *Behavioral Sciences*, vol. 14, no. 10, p. 894, oct. 2024. doi: 10.3390/bs14100894.
- [11] "AI Regulation Costa Rica: Emerging Policy & Compliance." [En línea]. Disponible en: <https://digital.nemko.com/regulations/ai-regulation-costa-rica>
- [12] N. Mehrabi, F. Morstatter, N. Saxena, K. Lerman y A. Galstyan, "A Survey on Bias and Fairness in Machine Learning," *ACM Computing Surveys*, vol. 54, no. 6, pp. 115:1–115:35, jul. 2021. doi: 10.1145/3457607.
- [13] J. Yang, A. A. S. Soltan, D. W. Eyre, Y. Yang y D. A. Clifton, "An adversarial training framework for mitigating algorithmic biases in clinical machine learning," *npj Digital Medicine*, vol. 6, no. 1, p. 55, mar. 2023. doi: 10.1038/s41746-023-00805-y.
- [14] R. K. E. Bellamy et al., "AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias," *IBM Journal of Research and Development*, vol. 63, no. 4/5, pp. 4:1–4:15, jul. 2019. doi: 10.1147/JRD.2019.2942287.
- [15] R. Shokri, M. Stronati, C. Song y V. Shmatikov, "Membership Inference Attacks Against Machine Learning Models," en *2017 IEEE Symposium on Security and Privacy (SP)*, may. 2017, pp. 3–18. doi: 10.1109/SP.2017.41.
- [16] B. Yang, "Navigating Privacy Risks in Generative AI: Concerns, Challenges, and Potential Solutions," *Journal of Computing and Electronic Information Management*, vol. 20, no. 1, pp. 6–11, ene. 2026. doi: 10.54097/bx1ne091.
- [17] X. Ren, S. Yang, C. Zhao, J. McCann y Z. Xu, "Belt and Braces: When Federated Learning Meets Differential Privacy," *Communications of the ACM*, vol. 67, no. 12, pp. 66–77, nov. 2024. doi: 10.1145/3650028.
- [18] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," *Nature Machine Intelligence*, vol. 1, no. 5, pp. 206–215, may. 2019. doi: 10.1038/s42256-019-0048-x.
- [19] C. Molnar, *Interpretable Machine Learning: A Guide For Making Black Box Models Explainable*. Munich, Germany: Christoph Molnar, 2025.
- [20] B. Dykes, *Effective Data Storytelling: How to Drive Change with Data, Narrative, and Visuals*. Hoboken, NJ: John Wiley and Sons, Inc., 2020.
- [21] Q. V. Liao y J. W. Vaughan, "AI Transparency in the Age of LLMs: A Human-Centered Research Roadmap," *Harvard Data Science Review*, no. Special Issue 5, may. 2024. doi: 10.1162/99608f92.8036d03b.
- [22] U. Bhatt et al., "Uncertainty as a Form of Transparency: Measuring, Communicating, and Using Uncertainty," en *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21)*. New York, NY, USA: Association for Computing Machinery, jul. 2021, pp. 401–413. doi: 10.1145/3461702.3462571.
- [23] L. Dittmar, "What Does Transparency Really Mean in the Context of AI Governance?," *OCEG*. [En línea]. Disponible en: <https://www.oceg.org/what-does-transparency-really-mean-in-the-context-of-ai-governance/>
- [24] Q. Booker, C. Rebman Jr., H. Wimmer, S. Levkoff, L. Powell y J. Breese, "Data Analytics Position Description Analysis: Skills Review and Implications for Data Analytics Curricula," *ISEDJ*, vol. 22, no. 3, p. 76, jul. 2024.
- [25] R. Rodrigues y A. Dias, "Influence of soft skills on career development: Exploring performance appraisal as a mediating mechanism," *Problems and Perspectives in Management*, vol. 22, no. 4, pp. 14–24, oct. 2024. doi: 10.21511/ppm.22(4).2024.02.
- [26] C. A. M. Sutherland y A. W. Young, "Understanding trait impressions from faces," *British Journal of Psychology*, vol. 113, no. 4, pp. 1056–1078, nov. 2022. doi: 10.1111/bjop.12583.
- [27] J. K. Burgoon, X. Wang, X. Chen, S. J. Pentland y N. E. Dunbar, "Nonverbal Behaviors 'Speak' Relational Messages of Dominance, Trust, and Composure," *Frontiers in Psychology*, vol. 12, ene. 2021. doi: 10.3389/fpsyg.2021.624177.
- [28] Ç. Doğru, "A Meta-Analysis of the Relationships Between Emotional Intelligence and Employee Outcomes," *Frontiers in Psychology*, vol. 13, abr. 2022. doi: 10.3389/fpsyg.2022.611348.
- [29] N. Türker y E. Üngüren, "Digital footprints and recruitment: an experimental study on the impact of social media content on hiring decisions," *Frontiers in Psychology*, vol. 16, oct. 2025. doi: 10.3389/fpsyg.2025.1693850.

30
AÑOS

CPIC
COLEGIO DE PROFESIONALES EN
INFORMÁTICA Y COMPUTACIÓN