TRABAJO DE GRADO
Opción Seminario-Diplomado.

Modelo de indicadores para evaluar la eficiencia, calidad y capacidad de respuesta ante incidentes de una mesa de ayuda de ciberseguridad tercerizada

Corporación Universitaria Remington
Facultad de Ingenierías
Ingeniería de Sistemas

Manuela Arango López
Jorge Mauricio Sepúlveda Castaño
Opción de Trabajo de grado Seminario
2026

Tabla de Contenidos

Resumen 3
Marco conceptual y contextual 5
1. El contexto: por qué se mide mal la ciberseguridad tercerizada 5
1.1. Qué se terceriza y cuál es la magnitud del problema. 5
1.2. El contexto colombiano. 7
2. Conceptos fundamentales 8
2.1. Mesa de ayuda, SOC y CSIRT: tres funciones que no se miden igual. 8
2.2. Medida, métrica, indicador y objetivo de nivel de servicio. 9
3. Marcos normativos de referencia 10
3.1. La taxonomía de medición del NIST. 10
3.2. Las normas ISO y la especificación de evaluación de CIS. 11
4. Estado del arte y vacío identificado 12
Desarrollo e implementación del aprendizaje 14
5. Método seguido 14
6. Indicadores de medición del outsourcing en ciberseguridad 15
6.1. Grupo A. Cumplimiento de niveles de servicio. 15
6.2. Grupo B. Métricas temporales del ciclo del incidente. 16
6.3. Grupo C. Cobertura. 17
6.4. Grupo D. Vulnerabilidades y parcheo. 18
6.5. Grupo E. Calidad de la detección. 19
6.6. Grupo F. Gestión de terceros, accesos y aseguramiento independiente. 20
6.7. Grupos G y H. Continuidad e indicadores económicos. 22
7. Medición de la eficiencia de la mesa de ayuda de ciberseguridad 23
7.1. Indicadores operativos y sus valores de referencia. 23
7.2. Indicadores propios del flujo de ciberseguridad. 24
7.3. Matriz de niveles de servicio por criticidad. 25
7.4. Tablero de control de cuatro perspectivas. 26
8. Por qué los indicadores se degradan y qué hacer al respecto 26
9. Caso de aplicación: seis meses de medición en una entidad vigilada 28
9.1. Escenario, alcance y advertencia sobre los datos. 29
9.2. Subconjunto de indicadores, metas, responsables y fuentes de evidencia. 30
9.3. Resultados de los seis meses. 30
9.4. Lectura del tablero y decisiones de mejora. 32
10. Tendencias que están modificando el objeto medido 34
Figuras y tablas 36
Conclusiones 43
Referencias 46

Resumen

Cada vez más organizaciones contratan a un proveedor externo para que opere su mesa de ayuda de ciberseguridad: el punto donde los usuarios reportan correos sospechosos, cuentas bloqueadas y comportamientos anómalos. Sin embargo, siguen evaluando ese servicio con los indicadores del soporte técnico convencional (cantidad de tickets, tiempo de atención y cumplimiento de los acuerdos de nivel de servicio), que miden velocidad y volumen, no reducción del riesgo. Una mesa puede cerrar todos sus casos a tiempo y aun así clasificar mal un incidente, de modo que el tablero registra cumplimiento mientras la exposición al riesgo permanece.

La participación de terceros en las brechas confirmadas pasó de cerca del 15 % en 2024 al 48 % en 2026, y cuando el ataque entra por la cadena de suministro la organización tarda 258 días en identificarlo y contenerlo, el ciclo más extenso entre los vectores estudiados.

Este informe técnico revisa la normativa y la literatura académica disponibles y construye a partir de ellas un modelo de indicadores agrupados en ocho conjuntos temáticos y organizados en cuatro perspectivas: operación, seguridad, servicio y gobierno del proveedor. Cada indicador se documenta con su fórmula, su valor de referencia cuando existe y, deliberadamente, cómo puede ser manipulado. Diez de ellos se ponen a prueba en un caso simulado de seis meses en una entidad financiera vigilada.

El hallazgo central es que la seguridad lograda no se puede observar: un día sin ataques es indistinguible de un día en que la prevención funcionó. Por eso todo contrato mide sustitutos, y esos sustitutos se degradan cuando se les asocia una consecuencia económica. Un acuerdo basado solo en penalidades llega a incentivar al proveedor a no

detectar incidentes, y el caso de aplicación muestra ese efecto con detalle. La respuesta del modelo tiene tres elementos inseparables: indicadores de segundo orden que revelan la manipulación de los primeros, verificación independiente del reporte y un contrato que premie la detección.

Se concluye que en Colombia las entidades vigiladas por la Superintendencia Financiera tienen la obligación regulatoria de contar con indicadores de ciberseguridad, pero no la especificación de cuáles. Este trabajo aporta ese contenido.

Palabras clave

Ciberseguridad tercerizada; indicadores de desempeño; mesa de ayuda; acuerdos de nivel de servicio; gestión de riesgo de terceros.

Marco conceptual y contextual

Este capítulo establece el terreno conceptual sobre el que se construye el modelo. Empieza por el contexto, esto es, qué se terceriza y con qué consecuencias medibles. Sigue con las definiciones que el trabajo usará de forma consistente y revisa los marcos normativos que dan sustento a la medición. Cierra identificando el vacío concreto que busca cubrir.

1. El contexto: por qué se mide mal la ciberseguridad tercerizada

1.1. Qué se terceriza y cuál es la magnitud del problema.

Tercerizar, según la norma ISO/IEC 20000-1:2018, es el acuerdo por el cual una organización externa ejecuta parte de una función de la organización contratante. La norma añade una nota de la que se desprende buena parte del planteamiento que sigue. La organización externa queda fuera del alcance del sistema de gestión del servicio, pero la función tercerizada permanece dentro (International Organization for Standardization e International Electrotechnical Commission [ISO e IEC], 2018). La ejecución se externaliza; la responsabilidad sobre la función permanece en la organización contratante. Dentro de ese conjunto, la mesa de ayuda de ciberseguridad ocupa una posición particular. Es el punto donde los reportes de los usuarios, sea un correo sospechoso, un bloqueo de cuenta o un comportamiento anómalo, se convierten, o no, en incidentes formalmente gestionados. Cuando esa función está en manos de un tercero, la organización pasa a depender de un proveedor para conocer su propio estado de seguridad. La guía NIST SP 800-61 Rev. 3 reconoce la modalidad y menciona expresamente la tercerización de un centro de operaciones de seguridad a un proveedor gestionado (Nelson et al., 2025).

El argumento a favor de medir mejor no es teórico. El Data Breach Investigations Report de Verizon mide la participación de terceros en las brechas confirmadas y registra tres años consecutivos de crecimiento, primero cerca del 15 %, luego 30 % y, en la edición de 2026, 48 % (Verizon, 2025, 2026). El propio informe advierte que se trata de una métrica de participación y no de causalidad.

El informe Cost of a Data Breach de IBM aporta la dimensión temporal. En 2026 el costo promedio global de una brecha alcanzó los 4,99 millones de dólares y el ciclo de identificación y contención promedió 247 días. Cuando el vector inicial fue el compromiso de la cadena de suministro, en cambio, ese ciclo se extendió a 258 días, el más largo de todos los vectores (IBM Security y Ponemon Institute, 2026). Ese es el dato más pertinente para este trabajo. El riesgo de terceros no solo aumenta la probabilidad del incidente; también deteriora la capacidad de detectarlo y contenerlo. La agencia europea de ciberseguridad añade que en estos ataques el 66 % de los atacantes se concentró en el código del proveedor (European Union Agency for Cybersecurity [ENISA], 2021), y sitúa el riesgo de cadena de suministro en el 10,6 % del total de amenazas registradas (ENISA, 2025).

Hay un dato adicional que interpela al objeto de este estudio. El informe de Verizon de 2026 documenta que los atacantes están suplantando a agentes de mesa de ayuda para obtener restablecimientos de credenciales y enrolamientos de segundo factor (Verizon, 2026). La mesa de ayuda, entonces, no es solo un mecanismo de defensa que conviene medir. Es también un vector de ataque, y esa doble condición debería reflejarse en los indicadores que se le exigen.

1.2. El contexto colombiano.

En Colombia, y para un sector concreto, el marco no es solamente una referencia de buenas prácticas. La Circular Básica Jurídica de la Superintendencia Financiera, en su Parte I, Título IV, Capítulo V, establece requerimientos mínimos para la gestión de la seguridad de la información y la ciberseguridad (Superintendencia Financiera de Colombia, 2020). Su alcance hay que delimitarlo bien. Estas exigencias obligan únicamente a las entidades vigiladas por esa Superintendencia, es decir, al sector financiero, asegurador y del mercado de valores, y no se extienden de manera automática a cualquier organización colombiana. Para las demás funcionan como referente sectorial de buena práctica. Tres numerales conforman el mandato regulatorio que sostiene este trabajo. El 3.9 obliga a incluir en los contratos con terceros críticos las medidas y obligaciones para cumplir las políticas de seguridad. El 3.10 exige verificar periódicamente ese cumplimiento. Y el 3.11 ordena contar con indicadores que midan la eficacia y la eficiencia de la gestión. La misma circular define al tercero crítico como aquel que, según los parámetros que fije la propia entidad, puede tener incidencia directa en la seguridad de su información.

A ese marco sectorial se suman dos leyes generales. La Ley 1273 de 2009 creó el bien jurídico de la protección de la información y tipificó el acceso abusivo a un sistema informático (Congreso de la República de Colombia, 2009), lo que obliga a delimitar contractualmente qué accesos autoriza el cliente al personal del proveedor. La Ley Estatutaria 1581 de 2012 exige conservar los datos personales bajo condiciones que impidan su adulteración, pérdida o acceso no autorizado (Congreso de la República de Colombia, 2012). Un proveedor que opera una mesa de ayuda accede a datos personales

de los usuarios del contratante, así que actúa como encargado del tratamiento, con deberes propios que el contrato debe recoger. En política pública, el documento CONPES 3995 de 2020 formuló la Política Nacional de Confianza y Seguridad Digital (Departamento Nacional de Planeación, 2020). Y la guía de gestión de incidentes del Ministerio de Tecnologías de la Información y las Comunicaciones (2021) aporta una fórmula de priorización y unos tiempos máximos de respuesta que sirven de referencia nacional.

2. Conceptos fundamentales

Conviene fijar antes tres términos. La guía NIST SP 800-161 Rev. 1 define proveedor como la organización o individuo que celebra un acuerdo con el adquirente para suministrar un producto o servicio, e incluye ahí a toda la cadena. Llama cuarta parte al proveedor del proveedor. Y llama riesgo de concentración a la exposición que aparece cuando varios terceros dependen de un mismo proveedor de cuarto nivel, circunstancia habitual en una mesa tercerizada, donde el proveedor trabaja con una plataforma de tickets, un motor de autenticación y una nube que él mismo contrata (Boyens et al., 2024). La misma guía nombra la dificultad de fondo, que es la asimetría de información entre adquirentes y proveedores. Los primeros se quedan sin visibilidad sobre cómo se prestan los servicios que contratan. Toda la discusión sobre indicadores que sigue puede leerse como un intento de compensar esa asimetría con evidencia verificable.

2.1. Mesa de ayuda, SOC y CSIRT: tres funciones que no se miden igual.

Uno de los errores más frecuentes en los contratos de servicios de seguridad consiste en aplicar a una función los indicadores de otra. El Foro de Equipos de Respuesta a Incidentes de Seguridad ofrece la distinción funcional más detallada entre las fuentes de acceso abierto consultadas (Forum of Incident Response and Security Teams [FIRST],

2019). Un equipo de respuesta a incidentes de seguridad informática (CSIRT) presta servicios a una comunidad definida para prevenir, detectar, manejar y responder a incidentes. Un centro de operaciones de seguridad (SOC) se concentra en el monitoreo y la detección de eventos. El criterio de separación es funcional. Los servicios obligatorios de un SOC son el monitoreo, la detección y el análisis de eventos, y los de un CSIRT son la recepción de reportes, el análisis del incidente y su coordinación (FIRST, s. f.). La mesa de ayuda no aparece en ese marco porque pertenece a otro dominio. Es el punto único de contacto de la gestión de servicios de TI con los usuarios, y su ubicación normativa está en las cláusulas de gestión de incidentes y solicitudes de ISO/IEC 20000-1:2018. En la práctica, una mesa de ayuda de ciberseguridad tercerizada ocupa una posición híbrida entre las tres figuras. Esa hibridez justifica construir un modelo de indicadores propio, porque ninguno de los conjuntos existentes cubre por sí solo el ámbito de la función. La Figura 1 representa el flujo a través de los tres niveles e indica en cada etapa qué indicadores se miden allí.

2.2. Medida, métrica, indicador y objetivo de nivel de servicio.

La guía NIST SP 800-55v1 introduce la terminología que este trabajo adopta (Schroeder et al., 2024a). La medición es el proceso de obtener valores cuantitativos. Las medidas son los valores objetivos que resultan de él. Y las métricas son medidas diseñadas para seguir el progreso respecto de un objetivo fijado, de modo que lo que separa a una medida de una métrica es la existencia de ese objetivo. La misma guía establece una consecuencia relevante para el outsourcing. Los resultados cualitativos, como alto, medio y bajo, y los semicuantitativos, como los niveles de madurez de uno a cinco, no constituyen

medidas. Un tablero de proveedor construido sobre semáforos o niveles de madurez no es medición, aunque se presente como tal.

En el plano contractual, la literatura sobre acuerdos de nivel de servicio de seguridad aporta tres términos más (Nicolazzo et al., 2024). El indicador de nivel de servicio es un parámetro medible, el objetivo es un valor umbral fijado sobre él y el acuerdo es el contrato que declara el nivel garantizado. Falta distinguir el indicador de desempeño del de riesgo. Siguiendo a ISACA (Tammineedi, 2018; Marquez, 2021), un indicador clave de riesgo muestra que la organización está sujeta a un riesgo que excede su apetito definido, de manera que no mide al proveedor: mide la exposición del cliente.

3. Marcos normativos de referencia

3.1. La taxonomía de medición del NIST.

La versión 2.0 del Marco de Ciberseguridad del NIST, de febrero de 2024 (National Institute of Standards and Technology, 2024), incorporó la función Gobernar. Su categoría de cadena de suministro exige que los proveedores sean conocidos y priorizados por criticidad, que los requisitos de ciberseguridad se integren en los contratos, que los riesgos del proveedor se monitoreen durante toda la relación y que los proveedores relevantes participen en la respuesta ante incidentes. El nombre de la función es significativo. Se llama gobernar porque la responsabilidad de gobierno no se transfiere.

En diciembre de 2024 el NIST reemplazó su guía de medición de seguridad de la información, vigente desde 2008 (Chew et al., 2008), por dos volúmenes. El primero se ocupa de identificar y seleccionar medidas y el segundo de desarrollar un programa de medición (Schroeder et al., 2024a, 2024b). Conviene precisarlo porque es fuente habitual de error, y es que no existe una revisión 2 de la SP 800-55. Su aporte central, y la pieza que

estructura todo este trabajo, es la taxonomía de cuatro tipos de medida. Las de implementación demuestran el progreso de controles específicos. Las de eficacia evalúan qué tan bien funcionan los controles ya implementados. Las de eficiencia examinan la velocidad a la que producen retroalimentación útil. Y las de impacto se refieren a resultados de misión, e incluyen expresamente las penalizaciones regulatorias y contractuales. Esa separación entre eficacia y eficiencia tiene una consecuencia decisiva, y es que el acuerdo de nivel de servicio de un proveedor de ciberseguridad resulta ser, predominantemente, un instrumento de medición de eficiencia y no de eficacia. El volumen 1 exige además documentar cada medida con identificador, meta, alcance, fórmula, valor objetivo, evidencia, responsables y fuente de datos, y la ficha del capítulo siguiente instancia ese requisito. Del volumen 2 interesa el continuo de madurez. Un programa nuevo debe empezar midiendo cobertura e implementación, que son verificables y difíciles de manipular, y dejar cronogramado en el contrato el desplazamiento hacia eficacia e impacto.

Queda una advertencia de método sobre la respuesta a incidentes. En abril de 2025 el NIST publicó la revisión 3 de su guía en la materia, que reemplazó y retiró la revisión 2 de 2012 y eliminó de paso la orientación operativa sobre métricas que esta contenía (Cichonski et al., 2012; Nelson et al., 2025). Por eso este trabajo recurre a la revisión 2, retirada pero conceptualmente vigente, para la taxonomía de métricas de respuesta, y advierte su estado cada vez que la cita.

3.2. Las normas ISO y la especificación de evaluación de CIS.

El Anexo A de ISO/IEC 27001:2022 (ISO e IEC, 2022c) dedica cinco controles consecutivos, del 5.19 al 5.23, a las relaciones con proveedores, desarrollados en ISO/IEC

27002:2022 (ISO e IEC, 2022b). La agencia europea de ciberseguridad identifica ese bloque como el estándar de referencia en la materia (ENISA, 2023). La serie ISO/IEC 27036 desarrolla el mismo dominio con más detalle, y de sus cuatro partes la segunda es la pertinente aquí, porque fija los requisitos para gestionar la relación durante todo su ciclo de vida (ISO e IEC, 2021, 2022a).

Ninguna de estas normas proporciona, sin embargo, fórmulas de cálculo. El vacío lo cubre la Especificación de Evaluación de Controles del Center for Internet Security, en su versión 8.1 (Center for Internet Security, 2024). Es la única fuente pública consultada que define fórmulas explícitas control por control, con una estructura de entradas, operaciones, medidas y métricas. Varias de las fichas del capítulo siguiente toman de allí su fórmula, y así se indica en cada caso.

4. Estado del arte y vacío identificado

La revisión realizada permite afirmar que existe literatura abundante sobre tres dominios contiguos, que son la medición de la seguridad de la información, la gestión del riesgo de terceros y el desempeño de las mesas de servicio. Sobre su intersección hay muy poca. El hallazgo más significativo es que la guía de referencia mundial en medición de seguridad, la SP 800-55 del NIST, no trata la medición de proveedores, cadena de suministro ni tercerización en ninguno de sus dos volúmenes (Schroeder et al., 2024a, 2024b). El volumen 1 remite ese dominio a la guía de riesgo de cadena de suministro y el volumen 2 asume que el programa de medición es propiedad de quien lo ejecuta. El estándar de referencia en medición de seguridad no cubre la medición del proveedor que presta el servicio.

En el terreno de las operaciones de seguridad el diagnóstico es igual de claro. El informe de madurez del modelo SOC-CMM de 2026 reporta que el 17 % de los centros participantes no cuenta con un programa de métricas, que el 33 % lo tiene implementado y optimizado y que el dominio de reporte apenas alcanza un puntaje de madurez de 1,6 (Van Os, 2026). La encuesta del SANS Institute de 2025 lo completa con un dato pertinente para este análisis, y es que el 69 % de los centros todavía depende de procesos manuales o mayoritariamente manuales para reportar sus métricas (Crowley, 2025). La medición está formalmente establecida, pero su producción sigue siendo manual. El mismo informe de madurez aporta además un hallazgo contraintuitivo, y es que los centros operados por proveedores gestionados alcanzan puntajes más altos que los internos. Eso desplaza la pregunta hacia donde este trabajo la sitúa, que no es si conviene tercerizar sino cómo medir lo tercerizado.

Finalmente, la revisión de literatura arbitrada en español, en repositorios iberoamericanos, no localizó trabajos que aborden específicamente indicadores de medición del outsourcing en ciberseguridad. Existe producción hispanoamericana sobre ciberseguridad en general y sobre tercerización de tecnologías de la información en general, como el modelo de gobierno para universidades públicas colombianas que proponen Sanguino Reyes y Cuesta Quintero (2023). Pero la intersección de ambos temas está esencialmente vacía. Este trabajo se propone contribuir a llenarla desde el contexto colombiano.

Desarrollo e implementación del aprendizaje

Este capítulo construye el modelo. Declara el método seguido, propone el conjunto de indicadores organizado en grupos temáticos, los integra en un tablero de cuatro perspectivas y en una matriz de niveles de servicio por criticidad, analiza por qué los indicadores se degradan al convertirse en objetivos contractuales, lleva ese análisis a un caso de aplicación simulado de seis meses y cierra con las tendencias que están modificando el objeto medido.

5. Método seguido

El trabajo es de naturaleza documental y avanzó en cuatro pasos. Primero delimité el objeto, que es una mesa de ayuda de ciberseguridad operada por un tercero bajo contrato, con funciones de recepción, clasificación, atención y escalamiento de eventos, solicitudes e incidentes de seguridad. Segundo, reuní material de tres tipos de fuente, que fueron normas y guías técnicas de organismos reconocidos, literatura revisada por pares e informes primarios de firmas con metodología publicada. Descarté los materiales de mercadeo de fabricantes y los agregadores de contenido por carecer de método declarado. Tercero, verifiqué cada dato y cada definición contra la fuente original y no contra resúmenes de terceros. El informe declara de forma explícita cuándo un dato no pudo verificarse en fuente primaria, cuándo una cifra proviene de un estudio cualitativo y cuándo una fórmula es construcción propia. Esa declaración no es un formalismo. La literatura gris sobre indicadores de seguridad contiene numerosas cifras que circulan sin fuente identificable, y distinguirlas de las que sí la tienen constituyó una parte sustancial del trabajo.

Cuarto, documenté cada indicador en una ficha con la estructura que exige la guía NIST SP 800-55v1, es decir, nombre, qué mide, fórmula, unidad, valor de referencia y responsable de la medición. Le añadí un campo que la guía no pide, la limitación conocida, porque un catálogo de indicadores que no declara cómo puede manipularse cada uno es un catálogo incompleto. La Tabla 1 presenta el resultado de esa clasificación para los indicadores del modelo. Todas las figuras y tablas del informe se agrupan en el capítulo que sigue al desarrollo, y cada una se menciona en el apartado que la utiliza.

6. Indicadores de medición del outsourcing en ciberseguridad

Los indicadores se agrupan en ocho conjuntos. Cada grupo se presenta con su fundamento, su clasificación según la taxonomía NIST y las limitaciones que condicionan su uso contractual.

6.1. Grupo A. Cumplimiento de niveles de servicio.

El indicador más extendido en los contratos de servicios de seguridad es el porcentaje de cumplimiento del acuerdo de nivel de servicio, calculado como el cociente entre los eventos de una severidad dada atendidos dentro del objetivo pactado y el total de eventos de esa severidad en el periodo. Es una medida de eficiencia en el sentido de NIST, porque mide oportunidad y no resultado.

No existe un valor de referencia externo para este indicador. Cualquier cifra que circule como estándar de la industria proviene de material comercial, porque el objetivo pactado es fruto de una negociación y no de una medición. El indicador arrastra además tres debilidades documentadas. Es manipulable mediante la reclasificación de severidad. Es vulnerable a las pausas del reloj por espera del cliente. Y mide la respuesta, no la eficacia, lo que puede producir un tablero favorable acompañado de un resultado de

seguridad deficiente. El diseño que se recomienda tiene tres piezas. Fijar la severidad con criterios objetivos auditables por el cliente. Medir la distribución mensual de severidades asignadas, porque una migración sostenida hacia severidades bajas es la señal temprana de manipulación. Y publicar el tiempo total de reloj pausado como dato independiente.

6.2. Grupo B. Métricas temporales del ciclo del incidente.

Las siglas de tiempo de uso corriente en la industria, MTTD para la detección, MTTA para la toma de la alerta y MTTR para lo que sigue, carecen de definición normativa. La revisión realizada no las encontró en ISO/IEC 27001, 27002, 27004 ni 27035, ni en NIST SP 800-55 o 800-61. Se propone por eso una definición operacional anclada en seis instantes del ciclo de respuesta, que son el compromiso inicial, la generación de la alerta válida, la toma de la alerta por un analista humano, la contención de la propagación, la restauración del servicio y la remediación de la causa raíz. Los intervalos entre ellos definen los indicadores temporales, según se detalla en la Figura 2 y en la Tabla 2. El problema se concentra en la R final de MTTR, que se usa indistintamente para responder, restaurar, recuperar, remediar y resolver. Este trabajo evita la sigla y emplea dos indicadores con nombres distintos, el tiempo medio de restauración del servicio y el tiempo medio de remediación de la causa raíz.

Dos advertencias condicionan su uso. El tiempo medio de detección depende de un instante, el compromiso inicial, que por definición no fue observado cuando ocurrió y solo se reconstruye para los incidentes que llegaron a detectarse. Un proveedor con detección deficiente reporta un tiempo bajo porque solo mide los incidentes fáciles. El indicador tiene, pues, sesgo de supervivencia y no debe usarse aislado. El tiempo medio de toma de la alerta admite manipulación sencilla, porque un reconocimiento automático o un script

que abre el ticket lo satisfacen sin que haya habido trabajo analítico, y debe medirse junto con el tiempo hasta el primer análisis sustantivo.

El contrapeso metodológico frente a ese sesgo es el porcentaje de detección interna, la proporción de incidentes descubiertos por la propia organización o su proveedor frente al total de confirmados. No puede mejorarse ignorando los incidentes difíciles. Dejar de detectar lo empeora, porque el incidente terminará apareciendo por vía externa. El informe M-Trends de 2026 reporta detección interna en el 52 % de los casos, frente al 43 % del año anterior (Mandiant, 2026).

Un indicador adicional, específico de la relación contractual, es el tiempo de notificación del incidente al cliente. Su calibración encierra el error contractual más frecuente. Cuando el cliente está sujeto a un plazo regulatorio de notificación, el objetivo pactado con el proveedor debe ser estrictamente menor que ese plazo, con margen para el triaje y la decisión jurídica. Pactar con el proveedor el mismo plazo que la ley concede al cliente consume el término legal completo y deja al cliente en incumplimiento.

6.3. Grupo C. Cobertura.

Los indicadores de cobertura son medidas de implementación y comparten una misma estructura, el cociente entre los elementos correctamente configurados y los que deberían estarlo. La Especificación de Evaluación de Controles del Center for Internet Security (2024) la define para la autenticación multifactor, desagregada en cuentas de aplicaciones expuestas externamente, acceso remoto y cuentas administrativas, y también para la recolección de registros de auditoría, con indicadores complementarios de retención, sincronización horaria y centralización. La sincronización horaria merece un comentario. Suele tratarse como un detalle menor cuando en realidad es la condición de

posibilidad del grupo anterior, porque sin relojes sincronizados entre las fuentes de registro los intervalos del ciclo del incidente no son calculables de manera defendible ante una auditoría.

Su debilidad estructural está en el denominador, que suele definir el propio proveedor. Excluir activos del inventario infla el indicador sin mejorar la seguridad, así que lo que debe auditarse es el inventario y no el porcentaje. La cobertura de autenticación multifactor arrastra además otra limitación, y es que no distingue la fuerza del segundo factor. Un mensaje de texto y una llave física cuentan igual, pese a resistir el phishing de forma muy distinta.

Un caso particular es la cobertura de técnicas de la matriz MITRE ATT&CK, para la que no existe fórmula normativa. La literatura arbitrada propone clasificar cada técnica como cubierta, brecha con capacidad técnica disponible o brecha sin fuente de registro (Chamkar et al., 2024). La limitación es severa. El denominador depende del perfil de amenaza y de la superficie tecnológica de cada organización, de modo que comparar porcentajes de cobertura entre organizaciones es metodológicamente inválido. La regla contractual que se propone es exigir que la cobertura declarada se valide mediante emulación de adversario y se reporte como detectado, alertado o bloqueado, nunca como regla existente.

6.4. Grupo D. Vulnerabilidades y parcheo.

Este grupo cuenta con las fórmulas mejor documentadas del catálogo, todas provenientes de la especificación del Center for Internet Security (2024): efectividad de actualización, cobertura de automatización, cumplimiento de escaneo, tasa de remediación

y cobertura de escaneo autenticado, con requisitos de temporalidad de escaneo mensual para activos externos y trimestral para internos.

Tres limitaciones acompañan su lectura. La primera es que la fórmula de efectividad de actualización incluye en el numerador las excepciones autorizadas. Un proveedor puede entonces alcanzar el cien por ciento convirtiendo activos desactualizados en excepciones, así que todo contrato debe medir por separado la tasa de excepciones y su tendencia. Una tasa creciente de excepciones acompañada de efectividad constante constituye un indicio consistente de manipulación. La segunda es que la tasa de remediación depende del intervalo entre escaneos, de modo que escanear con menos frecuencia la mejora artificialmente, y debe reportarse siempre junto con el número de días entre escaneos. La tercera es la más importante. La fórmula mide la desaparición de la vulnerabilidad del reporte de escaneo, no su remediación, y un activo apagado, retirado del alcance o escaneado sin autenticación también desaparece. De ahí que la cobertura de escaneo autenticado sea el indicador de control que impide falsear el resto del grupo.

Como referencia del estado real de la industria, y no como objetivo a alcanzar, el informe de Verizon documenta que la remediación completa de vulnerabilidades en dispositivos de borde y redes privadas virtuales tomó una mediana de 32 días y que solo alrededor del 54 % fueron remediadas por completo durante el año, y que en la edición de 2026 apenas el 26 % de las vulnerabilidades críticas fue totalmente remediado, en descenso desde el 38 % (Verizon, 2025, 2026).

6.5. Grupo E. Calidad de la detección.

La tasa de falsos positivos es el indicador que con mayor frecuencia se interpreta de forma incorrecta, y su análisis exige precaución con las fuentes. Un estudio cualitativo

presentado en el simposio de seguridad de USENIX recogió testimonios de analistas según los cuales uno de cada cien casos investigados corresponde a una amenaza real (Alahmadi et al., 2022). Conviene precisar que son percepciones y no una tasa medida. El mismo estudio aporta, en cambio, una distinción verificada y con consecuencias operativas directas. No es lo mismo una falsa alarma, que es un fallo genuino de detección sin evento real detrás, que un disparo benigno, que es una alarma que coincide con la firma pero corresponde a actividad legítima del negocio. La tasa mezcla entonces dos fenómenos con responsables distintos, porque el origen de los disparos benignos está en que el cliente no comunicó qué actividad es legítima en su entorno.

La limitación decisiva opera en dirección contraria a la intuitiva. Un proveedor reduce su tasa de falsos positivos elevando umbrales y escalando menos, lo que aumenta los falsos negativos, que son invisibles. Penalizar contractualmente la tasa de falsos positivos sin medir a la vez los falsos negativos empeora la seguridad de la organización. Constituye una manifestación característica de la ley de Goodhart dentro de un acuerdo de nivel de servicio. La tasa de falsos negativos, por su parte, mide una cantidad inobservable por definición, y su único estimador con denominador conocido es la emulación de adversario, esto es, la proporción de técnicas ejecutadas en el ejercicio que no fueron detectadas. La regla contractual es que los escenarios los defina el cliente y que el proveedor no conozca el calendario.

6.6. Grupo F. Gestión de terceros, accesos y aseguramiento independiente.

Dos indicadores componen este grupo. El primero es el porcentaje de proveedores críticos con evaluación de seguridad vigente, que responde al criterio de servicios de confianza del instituto americano de contadores públicos sobre evaluación periódica de

proveedores y socios de negocio (American Institute of Certified Public Accountants, 2022a) y a la cláusula de medición de la parte 2 de ISO/IEC 27036 (ISO e IEC, 2022a). Su limitación es doble. La criticidad de los subproveedores la clasifica el propio evaluado, de modo que reclasificar a un tercero como no crítico lo saca del denominador. Y una evaluación vigente no equivale a una satisfactoria, por lo que debe acompañarse del porcentaje de proveedores críticos con hallazgos abiertos vencidos.

El segundo es la revocación oportuna de accesos de terceros, donde la limitación es de fondo. El indicador depende de que el proveedor notifique las desvinculaciones de su propio personal, y ese es un evento que el cliente no observa. La recomendación es medir por conciliación independiente, comparando la nómina declarada por el proveedor contra las identidades activas en los sistemas del cliente. Conviene atender también la brecha entre cuenta deshabilitada y acceso efectivamente revocado, donde persisten sesiones activas, tokens, llaves de acceso y cuentas de servicio compartidas.

Un tercer indicador cierra el grupo y responde a una pregunta que el modelo no puede eludir. Quién verifica lo que el proveedor reporta. La respuesta contractual habitual es delegar esa verificación en un tercero acreditado, y los tres mecanismos disponibles tienen alcances que conviene no confundir. El informe SOC 2 de tipo II evalúa los controles de una organización de servicio contra los criterios de servicios de confianza y verifica su eficacia operativa a lo largo de un periodo (American Institute of Certified Public Accountants, 2022a). Sus límites, sin embargo, son precisos. El alcance lo define la organización de servicio y no el cliente, los criterios de descripción permiten dejar fuera a las suborganizaciones y los controles complementarios trasladan responsabilidades al

cliente (American Institute of Certified Public Accountants, 2022b). Un proveedor puede, por tanto, exhibir un informe sin salvedades con su cadena crítica excluida. La norma internacional de trabajos de aseguramiento 3402 (International Auditing and Assurance Standards Board, 2011) produce informes de estructura similar, pero su alcance normativo es el control interno sobre el reporte financiero y no la ciberseguridad; aceptarla como evidencia de postura de seguridad es un error frecuente en los contratos. Y la certificación ISO/IEC 27001, emitida por organismos acreditados bajo ISO/IEC 27006-1:2024 (ISO e IEC, 2024), acredita la conformidad de un sistema de gestión dentro de un alcance declarado, no la eficacia de la seguridad. Es un documento de una página, sin hallazgos ni excepciones. De ahí el indicador que el modelo denomina índice de aseguramiento independiente vigente y pertinente, la proporción de proveedores críticos con un informe de aseguramiento vigente cuyo alcance cubre efectivamente el servicio contratado y sin excepciones abiertas relevantes. La triple condición no constituye un refinamiento accesorio. Contabilizar cuántos proveedores disponen de un informe SOC 2, sin examinar su alcance ni sus excepciones, produce un indicador sin capacidad diagnóstica.

6.7. Grupos G y H. Continuidad e indicadores económicos.

Los indicadores de continuidad, que son los ejercicios exitosos y el cumplimiento de los objetivos de tiempo y de punto de recuperación, se apoyan en la guía de planeación de contingencia del NIST (Swanson et al., 2010). Esa guía define el objetivo de tiempo de recuperación como el máximo que un recurso puede permanecer indisponible antes de un impacto inaceptable, y el de punto de recuperación como el momento previo a la interrupción hasta el cual pueden recuperarse los datos. Ambos están acotados por el tiempo máximo tolerable de interrupción, que lo decide el negocio. De ahí se sigue algo

que conviene subrayar: pactar un objetivo de recuperación sin haber determinado antes ese máximo carece de fundamento.

La limitación práctica es que se miden casi siempre en pruebas planificadas y no en desastres reales, de modo que evalúan el ensayo y no la función. A eso se añade que el objetivo de punto de recuperación se cumple formalmente con la sola existencia del respaldo, aunque sea irrecuperable. Debe medirse restauración verificada. Los indicadores económicos del Grupo H son el costo por incidente, el costo por ticket y las penalizaciones contractuales efectivamente aplicadas, las únicas medidas de impacto del modelo. Queda fuera de él, deliberadamente, el retorno de la inversión en seguridad (ENISA, 2012; Gordon y Loeb, 2002). No es un indicador de desempeño del proveedor sino del retorno de una inversión, y su lugar es la justificación previa de la decisión de tercerizar, no el contrato.

7. Medición de la eficiencia de la mesa de ayuda de ciberseguridad

Los indicadores del apartado anterior aplican a cualquier servicio de seguridad tercerizado. Este se concentra en los específicos de la mesa de ayuda, que deben leerse en dos capas. Una es la operativa, heredada de la gestión de servicios de TI. La otra es la de seguridad, que es la que distingue a esta función de una mesa convencional.

7.1. Indicadores operativos y sus valores de referencia.

La medición del desempeño de mesas de servicio tiene una tradición más larga y datos de referencia publicados. El trabajo de referencia propone un cuadro de mando de siete indicadores con pesos relativos, encabezado por el costo por contacto y la satisfacción del cliente con un 25 % cada uno, seguidos de la satisfacción del agente con un 20 %, la resolución en primer contacto con un 15 %, la utilización del agente con un 10 % y la velocidad media de respuesta con un 5 % (Rumburg y Zbikowski, 2012). De esa misma

fuente proviene la serie de costos por nivel de escalamiento. El costo por incidente crece de forma pronunciada conforme sube el nivel que resuelve el caso, con valores del orden de 22, 62, 85, 196 y 471 dólares. Lo que sustenta el argumento económico de la resolución en primer contacto, y buena parte del caso de negocio de la tercerización, es la proporción entre niveles y no los valores absolutos.

Las series de referencia que circulan en las publicaciones de benchmarking del sector sitúan la resolución en primer contacto en torno al 74 %, la satisfacción del cliente en el 74,4 %, la utilización del agente en el 48,4 % y el tiempo medio de resolución en 8,85 horas hábiles, con una rotación cercana al 40 %. Aquí corresponde una declaración, conforme al método del apartado 5. Esos valores no pude verificarlos en un informe primario de acceso público, porque proceden de materiales difundidos de forma fragmentaria y sin metodología muestral publicada. Los recojo por su valor orientativo, pero no deben usarse como objetivos contractuales, porque corresponden a mesas de servicio generales y no de seguridad y provienen de una muestra autoseleccionada. La propia literatura lo confirma al publicar, desde la misma fuente, al menos tres rangos distintos para el umbral de utilización a partir del cual se dispara la rotación.

7.2. Indicadores propios del flujo de ciberseguridad.

Lo que distingue a una mesa de ayuda de ciberseguridad no es el volumen ni la velocidad, sino la consecuencia del error de clasificación. Una mesa que cierre tickets con rapidez pero clasifique mal un reporte de phishing, o que no escale una alerta crítica, aparenta eficiencia mientras incrementa el riesgo de la organización. Por eso el modelo incorpora seis indicadores que conectan la operación con el riesgo.

- Precisión de clasificación: tickets correctamente categorizados sobre tickets auditados. Exige una auditoría muestral periódica, porque no puede calcularse desde el sistema de tickets.
- Cumplimiento de escalamiento crítico: casos críticos escalados dentro del plazo sobre el total de casos críticos. Es el indicador con menor tolerancia al incumplimiento.
- Falsos negativos de escalamiento: incidentes confirmados que no fueron escalados oportunamente sobre el total de incidentes confirmados. Es el complemento indispensable del anterior.
- Tasa de reapertura: tickets reabiertos sobre tickets cerrados. Detecta cierres prematuros y contrapesa cualquier indicador de velocidad.
- Reincidencia: incidentes repetidos por causa, activo o vulnerabilidad sobre el total. Distingue la atención sintomática de la erradicación.
- Calidad de cierre: tickets con evidencia, causa y solución documentadas sobre tickets cerrados auditados. Es el indicador que hace posible el aprendizaje y la trazabilidad forense.

Una recomendación transversal: no debe emplearse un tiempo medio de resolución global. Comparar el desbloqueo de una cuenta con la contención de un cifrado por ransomware produce un promedio carente de significado interpretativo, de modo que la desagregación por severidad es una condición de validez del indicador y no una opción de presentación.

7.3. Matriz de niveles de servicio por criticidad.

La Tabla 3 propone una matriz de niveles de servicio diferenciada por prioridad. Los valores son una propuesta de diseño académico y deben ajustarse al contrato, al tamaño de la organización y a la regulación aplicable, así que no constituyen un estándar. Su utilidad reside menos en las cifras que en la exigencia que imponen. Para cada nivel debe definirse explícitamente qué se entiende por primera respuesta, escalamiento, contención,

resolución y cierre. Sin esas definiciones, un proveedor puede cumplir formalmente el acuerdo enviando una respuesta automática sin haber iniciado ninguna acción técnica significativa.

7.4. Tablero de control de cuatro perspectivas.

Los indicadores se integran en un tablero mensual de cuatro perspectivas, presentado en la Tabla 4 y representado en la Figura 3. Cada perspectiva responde a una pregunta distinta. Si la mesa responde y resuelve con eficiencia. Si contribuye a reducir el riesgo. Si el usuario recibe una atención oportuna y confiable. Y si el proveedor cumple, reporta con veracidad y mejora.

La lectura debe ser conjunta y no indicador por indicador. Si los casos de prioridad uno cumplen el tiempo de primera respuesta pero el tiempo de contención aumenta y la reapertura crece, la conclusión correcta no es que dos indicadores están en verde y dos en rojo, sino que la mesa es rápida para registrar casos e insuficiente para contenerlos de forma sostenible. Los indicadores adquieren significado en sus relaciones, no en sus valores individuales.

8. Por qué los indicadores se degradan y qué hacer al respecto

Este apartado sostiene el argumento central del trabajo. No es una advertencia general sobre el mal uso de las métricas, sino una cadena de cinco proposiciones con respaldo documentado. La Figura 4 la resume junto con las tres medidas de mitigación que se derivan de ella.

La primera proposición es que el resultado del esfuerzo preventivo en seguridad es estructuralmente inobservable. Ni la firma ni el proveedor pueden observar perfectamente el resultado del esfuerzo de prevención (Cezar et al., 2014). Es un problema de riesgo moral

distinto del habitual. No es que el principal no observe el esfuerzo del agente; es que ninguna de las dos partes observa el resultado, porque una prevención exitosa es indistinguible de la ausencia de ataque.

La segunda es que, en consecuencia, los contratos miden sustitutos observables como tiempos, coberturas y volúmenes, es decir, eficiencia y no eficacia. La tercera es que, al convertirse esos sustitutos en objetivos con consecuencia económica, dejan de ser buenas medidas. Es la ley de Goodhart (Hood y Piotrowska, 2021), formulada como que toda regularidad estadística observada comenzará a colapsar en cuanto se ejerza presión sobre ella con fines de control.

La cuarta precisa la naturaleza del fenómeno. La investigación sobre manipulación de cifras públicas caracteriza el gaming por tres rasgos. No es inequívocamente ilegal, porque explota zonas grises. No busca únicamente beneficio personal, sino que avanza intereses organizacionales. Y se concentra en los métodos de cálculo. Su mecanismo principal es la categorización creativa, que consiste en ubicar estratégicamente las unidades en categorías más favorables. Reclasificar un incidente de severidad uno a severidad dos cumple los tres rasgos de forma literal, lo que dificulta su cuestionamiento en sede contractual y hace que solo resulte detectable mediante indicadores de segundo orden.

La quinta es la más contraintuitiva y la de mayor consecuencia práctica. El mismo análisis demuestra formalmente que un esquema de penalidad exclusiva, es decir, tarifa fija más multa por brecha, desincentiva el esfuerzo de detección, porque descubrir una brecha dispara investigaciones que pueden terminar asignando responsabilidad al propio proveedor. En síntesis, un acuerdo basado únicamente en penalidades incentiva al

proveedor a no detectar. La predicción encuentra correlato empírico en el estudio de USENIX (Alahmadi et al., 2022), donde los analistas de proveedores gestionados reportan escalar un volumen elevado de alarmas hacia los clientes por precaución, trasladando el costo del triaje al cliente mientras mejoran su propio indicador.

La conclusión es que ningún conjunto de indicadores resuelve el problema por sí solo. La mitigación requiere tres elementos simultáneos. El primero es el diseño de incentivos, con la estructura que los autores proponen y que supera a las demás, que combina penalidad por brecha con recompensa por detección. El segundo es la verificación independiente del reporte mediante aseguramiento de tercero leído con criterio, en los términos del Grupo F. Y el tercero son los indicadores de segundo orden diseñados para detectar la manipulación de los de primer orden, como la distribución mensual de severidades asignadas, la tasa de excepciones de parcheo y su tendencia, la frecuencia real de escaneo, el porcentaje de detección interna y la tasa de falsos negativos estimada por emulación. La Tabla 5 desarrolla ese tercer elemento indicador por indicador, emparejando cada indicador de primer orden con la forma en que puede manipularse y con el indicador de segundo orden que lo revela.

9. Caso de aplicación: seis meses de medición en una entidad vigilada

Todo lo anterior es propuesta. Un catálogo de indicadores solo demuestra su utilidad cuando se lo pone a funcionar sobre datos, se lo lee con criterio y se derivan de esa lectura decisiones concretas. Este apartado hace ese ejercicio. Toma un subconjunto de diez indicadores del modelo, los sigue durante seis meses en un escenario construido para este trabajo y muestra qué habría visto el cliente en su tablero, qué habría pasado

inadvertido mirando solo el indicador contractual y qué decisiones se siguen de la lectura completa.

9.1. Escenario, alcance y advertencia sobre los datos.

El escenario es el de una entidad financiera colombiana de tamaño medio, vigilada por la Superintendencia Financiera, con cerca de 2.800 empleados y 41 oficinas. La atención de eventos, solicitudes e incidentes de seguridad está tercerizada en un proveedor que opera una mesa especializada 24 por 7, con dos niveles propios y escalamiento al equipo interno de respuesta para los casos de severidad mayor. El volumen ronda los 1.100 tickets de seguridad por mes. La elección de una entidad vigilada responde a un criterio explícito. Es el tipo de organización a la que le aplican las instrucciones de la Circular Básica Jurídica (Superintendencia Financiera de Colombia, 2020) sobre servicios prestados por terceros, de manera que el seguimiento y el acceso a la información del proveedor dejan de ser buenas prácticas y se vuelven exigencias supervisadas.

El punto de partida contractual es un acuerdo de nivel de servicio con un objetivo de tiempo de respuesta para los casos clasificados con severidad uno, en adelante P1, y con una penalidad económica por incumplirlo. Esa penalidad se introduce al comenzar el tercer mes de la serie, en una renegociación cuyo propósito declarado era mejorar la respuesta ante los casos críticos. El resto de los indicadores se mide desde el primer mes, sin consecuencia económica asociada.

Antes de presentar cualquier cifra corresponde una advertencia que considero parte del método y no un formalismo. Los datos del caso son simulados. Los construí para este ejercicio cuidando dos cosas: que fueran internamente consistentes, de modo que el número de casos P1 dividido por el total de tickets reproduzca exactamente la proporción reportada

en cada mes, y que la dirección de cada serie fuera coherente con lo que la literatura del apartado anterior predice. No constituyen evidencia empírica de que el fenómeno ocurra con esta magnitud, y sería un error citarlos como tal. Lo que el caso demuestra es más modesto y a la vez más útil. Demuestra que el conjunto de indicadores propuesto hace visible un patrón de degradación que el indicador contractual, por sí solo, oculta.

9.2. Subconjunto de indicadores, metas, responsables y fuentes de evidencia.

El subconjunto lo elegí con un criterio deliberado. Un solo indicador de primer orden, el que tiene consecuencia económica, acompañado de nueve capaces de revelar su manipulación o su desgaste. Es la estructura de mitigación del apartado anterior, reducida a un tamaño que una entidad puede sostener mes a mes. La Tabla 6 documenta cada indicador con su meta, el responsable de la medición y la fuente de evidencia, siguiendo la ficha que exige la guía de medición del NIST (Schroeder et al., 2024a).

Dos decisiones de diseño merecen justificarse. La primera es que el responsable de la medición no es el proveedor en todos los casos. Cinco de los diez indicadores los calcula la entidad a partir de datos crudos y no de reportes elaborados, y otros dos los reporta el proveedor con revalidación o muestreo independiente de la entidad, que es la única forma de sostener la verificación planteada en el Grupo F, y eso exige que el contrato garantice acceso a los registros y no solo a los tableros. La segunda es que la fuente de evidencia se declara para cada indicador. Un indicador sin fuente identificable no es medible en la práctica, aunque tenga fórmula, porque cuando llega el momento de discutir una cifra con el proveedor la conversación se decide sobre el registro que la respalda.

9.3. Resultados de los seis meses.

La Tabla 7 recoge la serie mensual de los ocho indicadores del subconjunto que se miden con esa periodicidad; los dos restantes, la rotación del personal asignado a la cuenta y el índice de aseguramiento independiente, se verifican con periodicidad anual y de vigencia, respectivamente, y por eso no forman serie mensual. La Figura 5 representa los tres indicadores que resumen el fenómeno, en dos paneles que comparten el eje de meses.

La lectura del indicador contractual es favorable y su cálculo no presenta objeciones. El cumplimiento del acuerdo en casos P1 pasa de 90,2 % en enero a 99,2 % en junio, la mejora se acelera justo después de introducirse la penalidad en el tercer mes y al cierre del semestre el proveedor supera el objetivo pactado. En un tablero que solo mostrara este indicador, el informe de gestión arrojaría un resultado favorable y la renegociación aparecería como una decisión acertada.

El resto de las series muestra un comportamiento divergente. La proporción de casos clasificados como P1 cae de 8,7 % a 3,1 %, a poco más de un tercio, mientras el volumen total de tickets se mantiene estable entre 1.062 y 1.143 por mes. No se registra menor actividad, sino un menor número de casos clasificados como críticos. Al mismo tiempo, el tiempo medio de contención de los casos que sí se clasifican como P1 sube de 5,1 a 9,4 horas, la tasa de reapertura pasa de 5,8 % a 13,9 %, la precisión de clasificación medida sobre muestra baja de 88 % a 74 %, el porcentaje de incidentes detectados por el propio proveedor cae de 56 % a 41 %, las excepciones de parcheo se triplican de 4,1 % a 12,3 % y la satisfacción del usuario desciende de 91 % a 82 %. Ninguna de esas siete series tiene consecuencia económica en el contrato.

El patrón es el que anticipa la cadena argumental del apartado 8. La medida se convirtió en objetivo con consecuencia y dejó de ser una buena medida (Hood y Piotrowska, 2021), y el mecanismo visible es la categorización creativa, ubicar los casos en la categoría más favorable en lugar de alterar el hecho. El efecto sobre la detección coincide con lo que el modelo formal predice para un esquema de penalidad exclusiva (Cezar et al., 2014). La caída del porcentaje de detección interna es el resultado de mayor relevancia de la serie, porque sugiere que el proveedor no solo reclasifica lo que encuentra, sino que ha dejado de buscar con la misma intensidad.

Ninguna de esas cifras, tomada de forma aislada, resulta concluyente. Una caída en la proporción de casos P1 puede reflejar una mejora real de la postura de seguridad, y es exactamente lo que el proveedor argumentó en el comité de seguimiento del cuarto mes. Lo que no admite esa explicación es la conjunción. Si hubiera menos incidentes críticos reales, el tiempo de contención de los que quedan no tendría por qué aumentar, la reapertura no debería crecer y la precisión de clasificación no debería deteriorarse. La fuerza probatoria está en el conjunto. Ese es el argumento de fondo a favor de medir en varias dimensiones a la vez.

9.4. Lectura del tablero y decisiones de mejora.

Con esa lectura sobre la mesa, la entidad ejecutó en el sexto mes una auditoría de conciliación. Tomó una muestra aleatoria de 40 tickets clasificados como P2 durante el cuatrimestre posterior a la renegociación y los revisó, con el equipo interno de respuesta, contra los criterios de severidad del anexo técnico. Once de los cuarenta cumplían criterios objetivos de P1, es decir, afectación de un sistema crítico, más de cincuenta usuarios impactados o indicio de compromiso de credenciales privilegiadas, y estaban registrados

como P2. El hallazgo convierte una sospecha estadística en un hecho verificable, y muestra por qué la verificación independiente del reporte no es un añadido opcional del modelo.

De la lectura del semestre se derivaron cuatro decisiones de mejora, que se enuncian aquí porque son el producto final del ejercicio de medición y no un apéndice suyo.

- Criterios objetivos de severidad en el anexo técnico. La clasificación dejó de depender del juicio del analista que recibe el caso y pasó a resolverse por una tabla de decisión con condiciones verificables: sistema afectado, número de usuarios impactados, tipo de dato involucrado e indicio de compromiso de credenciales. Es la respuesta directa a la categorización creativa, porque reduce la zona gris que el mecanismo explota.
- Auditoría muestral mensual. Cuarenta tickets al mes, elegidos al azar y revisados por la entidad, con la precisión de clasificación reportada como indicador propio.
- Rediseño del incentivo. La penalidad exclusiva se sustituyó por un esquema mixto de penalidad por incumplimiento más reconocimiento económico por detección propia verificada, en la línea de la estructura que Cezar et al. (2014) demuestran superior. El objetivo explícito es dejar de castigar el hallazgo.
- La distribución de severidades como indicador contractual. La proporción mensual de casos por severidad pasó a formar parte del reporte obligatorio, con umbrales de alerta ante variaciones superiores a dos puntos porcentuales respecto del promedio del trimestre previo.

El caso deja una lección de método que trasciende el escenario simulado. La entidad tenía el indicador correcto desde el primer mes. El cumplimiento del acuerdo en P1 medía exactamente lo que decía medir y su cálculo no presentaba errores metodológicos. El problema no estaba en el indicador sino en la ausencia de contrapesos, y el costo de esa ausencia fueron cuatro meses de degradación real presentada como mejora documentada.

Un sistema de medición reducido no equivale a un sistema de medición eficiente: la ausencia de contrapesos compromete la validez del conjunto.

10. Tendencias que están modificando el objeto medido

El modelo propuesto describe un servicio que está cambiando, y conviene cerrar el capítulo señalando en qué dirección.

La primera tendencia es la incorporación de inteligencia artificial al triaje, con categorización automática de tickets, enriquecimiento de alertas, agrupación de casos similares y consulta asistida de procedimientos de respuesta. Gartner (2026) la formula en términos críticos, al señalar que estas soluciones desestabilizan las normas operativas, y recomienda priorizar a las personas tanto como a la tecnología. Para el modelo tiene tres consecuencias. La decisión de cerrar, contener o escalar un incidente crítico debe conservar supervisión humana. El tablero debe incorporar un indicador de efectividad de la automatización que distinga los casos resueltos automáticamente de los cerrados sin resolución. Y si el proveedor introduce agentes automáticos en el flujo de atención, esos agentes son identidades no humanas con acceso a los sistemas del cliente, así que entran en el denominador de los indicadores de cobertura de autenticación y revocación de accesos del Grupo F.

La segunda es el desplazamiento desde métricas de volumen hacia métricas de resultado, que permite evaluar el contrato por reducción de exposición y resiliencia operativa antes que por cantidad de tickets cerrados. Esa orientación es coherente con la taxonomía de tipos de medida adoptada en este trabajo, aunque el continuo de madurez del NIST advierte que ese desplazamiento debe ser gradual y estar cronogramado.

La tercera es la integración creciente entre la mesa de ayuda, el centro de operaciones y el equipo de respuesta, apoyada en plataformas de orquestación, detección extendida y gestión de identidades. Su efecto sobre la medición es que las fronteras entre los tres conjuntos de indicadores se difuminan, lo que obliga a definir con precisión qué mide cada uno. La cuarta, muy ligada a la anterior, es la expansión de las identidades no humanas, es decir, servicios, interfaces de programación, dispositivos y aplicaciones, como superficie de riesgo relevante. Una mesa de ciberseguridad ya no gestiona únicamente cuentas de empleados, y los indicadores de cobertura de autenticación y de revocación de accesos deben ampliar su denominador en consecuencia.

La quinta, finalmente, es la fatiga por alertas y la rotación de analistas. La encuesta del SANS Institute de 2025 reporta que el 62 % de los profesionales de operaciones de seguridad considera que su organización no hace lo suficiente para retener al talento, y que apenas el 4 % permanece más de diez años en el puesto (Crowley, 2025). La continuidad del conocimiento sobre el entorno del cliente es, por tanto, frágil. Medir la rotación del personal asignado a la cuenta no es un indicador de recursos humanos, sino un indicador anticipado de degradación de la calidad del triaje, y debería figurar en la perspectiva de gobierno del proveedor.

Figuras y tablas

Este capítulo reúne las figuras y las tablas del modelo. Se presentan agrupadas aquí, y no intercaladas en el texto, para conservar la continuidad de la argumentación, conforme al formato de presentación de la Facultad. Cada figura y cada tabla se numera y se menciona en el apartado correspondiente del informe. Los títulos de las tablas se ubican sobre ellas y los pies de figura debajo, según lo exige ese mismo formato.

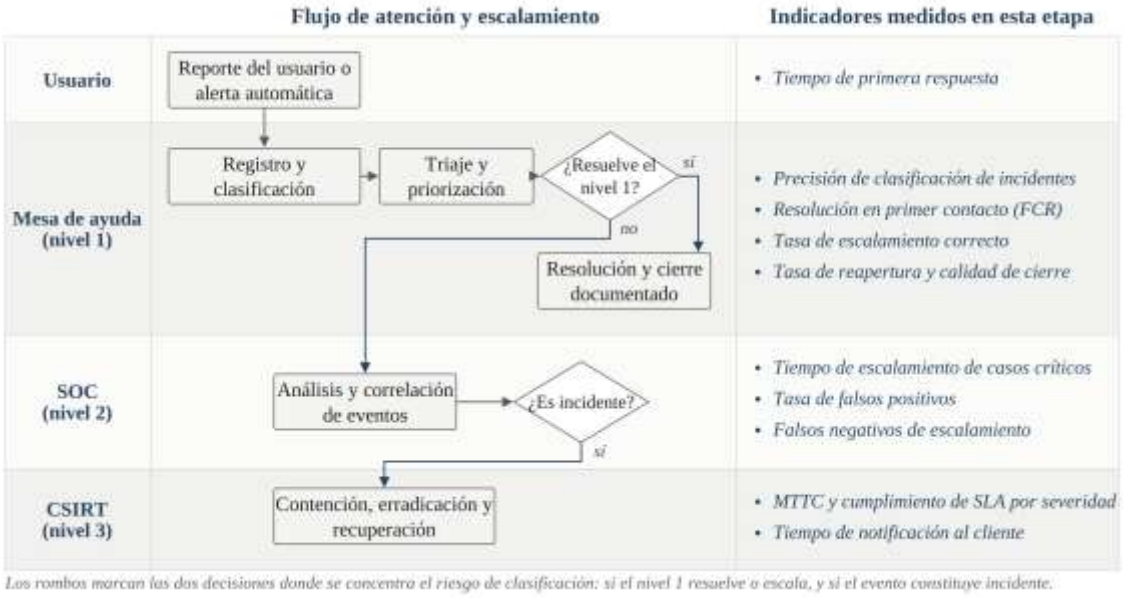


Figura 1. Flujo de atención y escalamiento de un evento de ciberseguridad a través de los tres niveles, con los indicadores que se miden en cada etapa. Elaboración propia a partir de la distinción funcional entre mesa de servicio, SOC y CSIRT establecida por FIRST (2019).

Tabla 1. Clasificación de los indicadores del modelo según la taxonomía NIST SP 800-55v1.

Tipo de medida	***Qué evalúa***	***Indicadores del modelo que pertenecen a este tipo***
Implementación	Progreso de controles específicos	Cobertura de autenticación multifactor; cobertura de registros y centralización; cobertura de monitoreo; proveedores críticos con evaluación vigente; índice de aseguramiento independiente; cobertura de técnicas ATT&CK; efectividad de actualización; cobertura de escaneo autenticado; ejercicios de continuidad realizados
Eficacia	Qué tan bien funcionan los controles	Precisión de clasificación; tasa de falsos positivos y de falsos negativos; porcentaje de detección interna; tasa de remediación por criticidad; reincidencia; tasa de reapertura; falsos negativos de escalamiento; calidad de

Tipo de medida	*Qué evalúa*	*Indicadores del modelo que pertenecen a este tipo*
		cierre; cumplimiento de los objetivos de tiempo y de punto de recuperación
Eficiencia	Oportunidad: velocidad de la retroalimentación útil	Cumplimiento del acuerdo de nivel de servicio; todos los indicadores temporales de la Tabla 2; resolución en primer contacto; revocación oportuna de accesos; cumplimiento de escalamiento crítico; tiempo de notificación al cliente
Impacto	Resultados de misión, incluidas penalizaciones	Costo por incidente; costo por ticket; penalizaciones contractuales aplicadas

Nota. Elaboración propia. La taxonomía de cuatro tipos de medida procede de NIST SP 800-55v1 (Schroeder et al., 2024a).

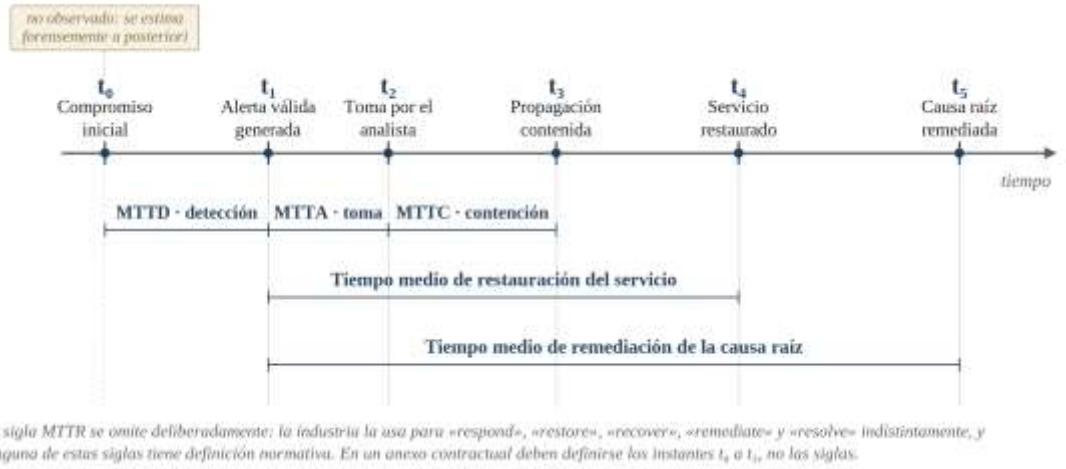


Figura 2. Línea de tiempo del incidente y definición operacional de los indicadores temporales. Elaboración propia sobre los hitos del ciclo de respuesta a incidentes de NIST SP 800-61 Rev. 2 (Cichonski et al., 2012), publicación retirada en abril de 2025 y utilizada aquí por su vigencia conceptual.

Tabla 2. Definición operacional de los indicadores temporales del ciclo del incidente.

Indicador	*Intervalo*	*Qué captura*	*Unidad*
Tiempo medio de detección	Alerta válida menos compromiso inicial	Capacidad de detección. Equivale al tiempo de permanencia del atacante.	Días
Tiempo medio de toma de la alerta	Toma humana menos alerta válida	Capacidad operativa del turno: dotación y cobertura horaria.	Minutos
Tiempo medio de contención	Contención menos toma humana	Rapidez para detener la propagación. No implica servicio restaurado.	Horas
Tiempo medio de restauración	Servicio restaurado menos alerta válida	Regreso del servicio a operación normal.	Horas o días
Tiempo medio de remediación	Causa raíz remediada menos alerta válida	Cierre de la causa raíz. Siempre igual o mayor que la restauración.	Días
Tiempo de notificación al cliente	Notificación menos calificación como incidente	Habilita al cliente a cumplir sus propios plazos legales.	Horas

Nota. Elaboración propia sobre los hitos del ciclo de respuesta a incidentes de NIST SP 800-61 Rev. 2 (Cichonski et al., 2012), publicación retirada en abril de 2025 y utilizada aquí por su vigencia conceptual. El tiempo de notificación al cliente se ancla en dos hitos adicionales de naturaleza contractual, la calificación formal como incidente y la notificación efectiva, que no forman parte de los seis instantes del ciclo técnico.

Tabla 3. Propuesta de niveles de servicio por criticidad para una mesa de ayuda de ciberseguridad tercerizada.

Prioridad	***Ejemplo***	***Primera respuesta***	***Escalamiento***	***Contención inicial***	***Actualización***
Crítica (P1)	Ransomware activo, cuenta privilegiada comprometida, fuga masiva de datos	15 minutos	30 minutos	4 horas	Cada 30 minutos
Alta (P2)	Phishing con credenciales comprometidas, malware confirmado en un equipo	30 minutos	1 hora	8 horas	Cada 2 horas
Media (P3)	Vulnerabilidad sin explotación confirmada, fallo de segundo factor	4 horas	8 horas	3 días hábiles	Diaria
Baja (P4)	Solicitud de información, orientación o ajuste no crítico	1 día hábil	Según necesidad	5 días hábiles	Según acuerdo

Nota. Elaboración propia. Los valores son una propuesta de diseño académico y no constituyen un estándar: deben ajustarse al contrato, al tamaño de la organización y a la regulación aplicable.

Tabla 4. Tablero de control de cuatro perspectivas.

Perspectiva	***Indicadores principales***	***Pregunta que responde***
Operación	Tickets recibidos, relación entre cerrados y recibidos, acumulado pendiente, tiempo medio de resolución por severidad, resolución en primer contacto, costo por ticket	¿La mesa responde y resuelve con eficiencia?
Seguridad	Tiempos de detección, toma y contención; precisión de clasificación; falsos negativos de escalamiento; porcentaje de detección interna; reincidencia; cobertura de monitoreo	¿La mesa contribuye a reducir el riesgo?
Servicio	Cumplimiento del acuerdo por prioridad, satisfacción del usuario, tasa de reapertura, tiempo de primera respuesta, calidad de cierre	¿El usuario recibe atención oportuna y confiable?
Gobierno del proveedor	Resultados de auditoría, índice de aseguramiento independiente, distribución mensual de severidades, tasa de excepciones de parcheo,	¿El proveedor cumple, reporta con veracidad y mejora?

Perspectiva	*Indicadores principales*	*Pregunta que responde*
	rotación del personal asignado, acciones correctivas cerradas	

Nota. Elaboración propia. La clasificación por tipo de medida sigue la taxonomía de NIST SP 800-55v1 (Schroeder et al., 2024a).

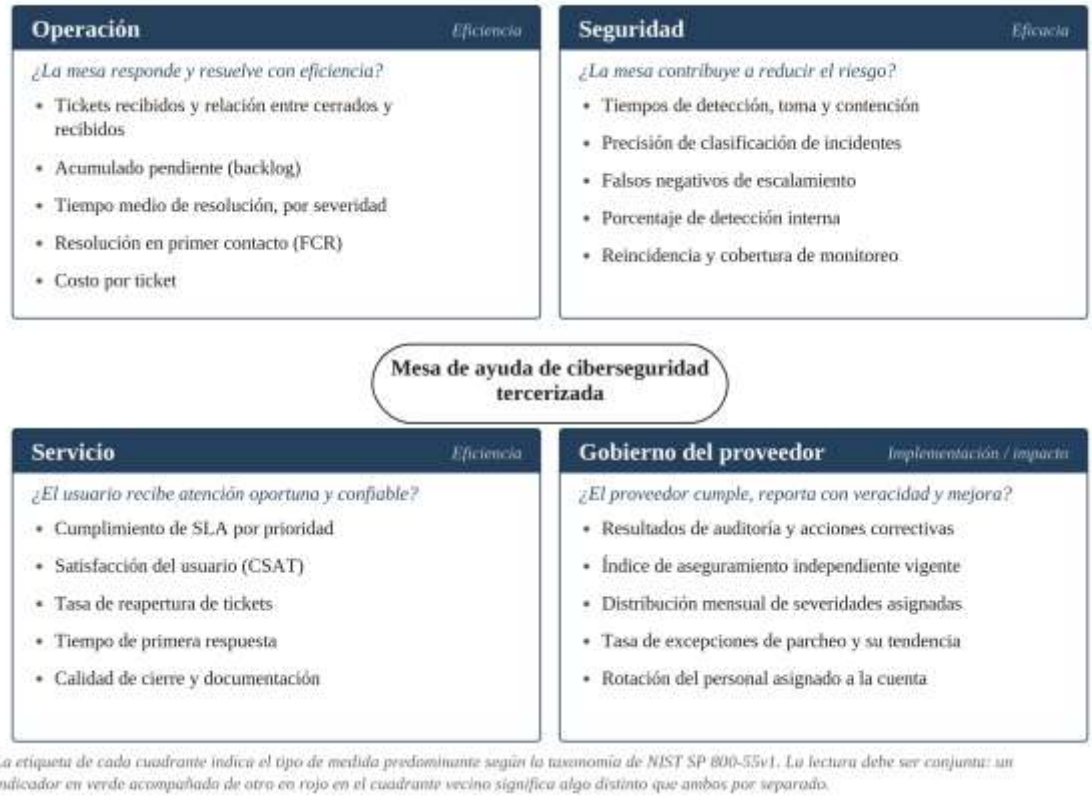


Figura 3. Tablero de control de cuatro perspectivas para una mesa de ayuda de ciberseguridad tercerizada. Elaboración propia; la clasificación por tipo de medida sigue la taxonomía de NIST SP 800-55v1 (Schroeder et al., 2024a).

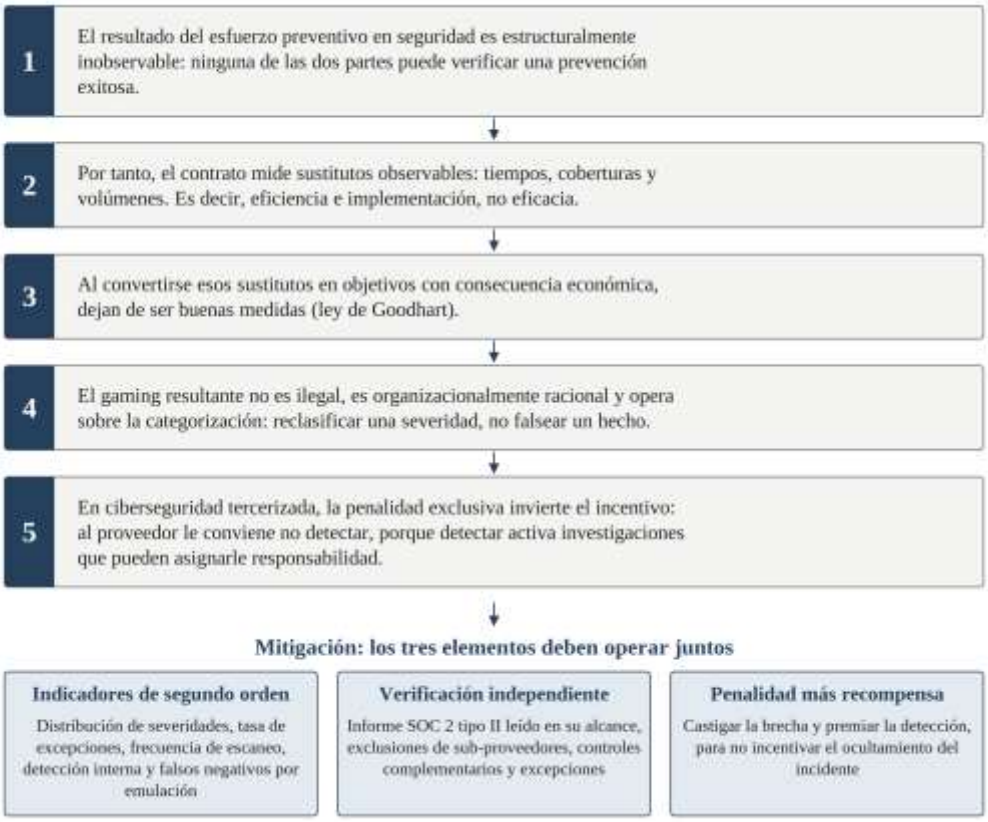


Figura 4. Cadena argumental de la degradación de los indicadores contractuales y las tres medidas de mitigación. Elaboración propia a partir de Cezar et al. (2014), Hood y Piotrowska (2021) y Alahmadi et al. (2022).

Tabla 5. Indicadores de segundo orden para detectar la manipulación de los de primer orden.

Indicador de primer orden	***Forma de manipulación***	***Indicador de segundo orden que la revela***
Cumplimiento del acuerdo de nivel de servicio	Reclasificación de severidad hacia abajo	Distribución mensual de severidades asignadas y tiempo total de reloj pausado
Tasa de falsos positivos	Elevación de umbrales de detección y menor escalamiento	Tasa de falsos negativos estimada por emulación de adversario definida por el cliente
Tiempo medio de detección	Se miden solo los incidentes fáciles de detectar	Porcentaje de detección interna frente a notificación externa
Efectividad de actualización de parches	Conversión de activos desactualizados en excepciones autorizadas	Tasa de excepciones sobre el total de activos y su tendencia
Tasa de remediación de vulnerabilidades	Reducción de la frecuencia de escaneo	Días entre escaneos y cobertura de escaneo autenticado
Cobertura de autenticación multifactor	Exclusión de activos del inventario	Auditoría independiente del inventario de activos y de cuentas

Indicador de primer orden	***Forma de manipulación***	***Indicador de segundo orden que la revela***
Tiempo medio de toma de la alerta	Reconocimiento automático sin análisis	Tiempo hasta el primer análisis sustantivo documentado
Revocación de accesos de terceros	El proveedor no notifica desvinculaciones de su personal	Conciliación entre la nómina declarada y las identidades activas en los sistemas

Nota. Elaboración propia a partir de Cezar et al. (2014), Hood y Piotrowska (2021) y Alahmadi et al. (2022).

Tabla 6. Subconjunto de indicadores del caso de aplicación, con metas, responsables de la medición y fuentes de evidencia.

Indicador	***Meta***	***Responsable de la medición***	***Fuente de evidencia***
Cumplimiento del acuerdo de nivel de servicio en casos P1	≥ 95 %	Proveedor, con revalidación de la entidad	Registros de la herramienta de gestión de tickets, exportados en crudo
Proporción de casos clasificados P1	Variación < 2 p. p. frente al promedio del trimestre anterior	Entidad	Conteo por severidad sobre el volcado mensual de tickets
Tiempo medio de contención de casos P1	≤ 6 h	Proveedor	Marcas de tiempo de contención en la bitácora del incidente
Precisión de clasificación	≥ 90 %	Entidad	Auditoría muestral de 40 tickets por mes contra los criterios del anexo técnico
Tasa de reapertura de tickets	≤ 6 %	Entidad	Tickets reabiertos dentro de los 7 días siguientes al cierre
Porcentaje de detección interna	≥ 60 %	Entidad	Origen del hallazgo registrado en cada incidente (proveedor, usuario o tercero)
Tasa de excepciones de parcheo	≤ 5 %	Proveedor	Inventario de excepciones vigentes con fecha de aprobación y de vencimiento
Satisfacción del usuario atendido	≥ 88 %	Proveedor, con muestreo independiente de la entidad	Encuesta posterior al cierre del ticket, con la base de respuestas disponible
Rotación del personal asignado a la cuenta	≤ 25 % anual	Proveedor	Nómina del equipo asignado, con fechas de ingreso y de retiro
Índice de aseguramiento independiente vigente y pertinente	100 % de los servicios críticos	Entidad	Informe SOC 2 de tipo II vigente, con revisión del alcance y de las excepciones

Nota. Elaboración propia. La estructura de la ficha sigue NIST SP 800-55v1 (Schroeder et al., 2024a). La abreviatura p. p. corresponde a puntos porcentuales.

Tabla 7. Resultados mensuales de los indicadores del caso de aplicación.

Indicador	***Ene***	***Feb***	***Mar***	***Abr***	***May***	***Jun***
Tickets de seguridad recibidos (n)	1.062	1.104	1.089	1.121	1.096	1.143
Casos clasificados P1 (n)	92	89	78	64	47	36
Proporción de casos P1 (%)	8,7	8,1	7,2	5,7	4,3	3,1
Cumplimiento del acuerdo de nivel de servicio en P1 (%)	90,2	91,0	93,6	96,9	98,3	99,2
Tiempo medio de contención de P1 (h)	5,1	5,4	5,6	6,8	8,1	9,4
Precisión de clasificación (%)	88	87	86	83	78	74
Tasa de reapertura (%)	5,8	6,1	6,4	9,2	11,7	13,9
Porcentaje de detección interna (%)	56	54	53	49	44	41
Tasa de excepciones de parcheo (%)	4,1	4,4	5,0	7,2	9,8	12,3
Satisfacción del usuario atendido (%)	91	90	90	88	85	82

Nota. Datos simulados, elaboración propia. La penalidad por incumplimiento del acuerdo de nivel de servicio en casos P1 se introduce al comenzar el tercer mes.

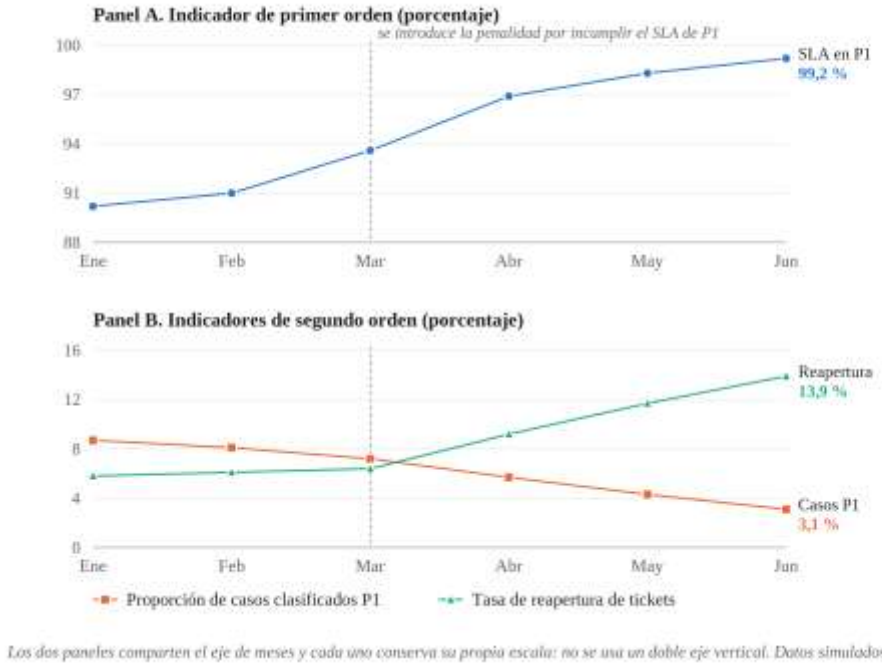


Figura 5. Divergencia entre el indicador contractual de primer orden y dos indicadores de segundo orden a lo largo de los seis meses del caso de aplicación. Los dos paneles comparten el eje de meses y cada uno conserva su propia escala. Elaboración propia sobre los datos simulados de la Tabla 7.

Conclusiones

La primera conclusión es que el problema de medir una mesa de ayuda de ciberseguridad tercerizada no es la falta de indicadores. Es que están mal clasificados. Existe una oferta abundante de métricas operativas heredadas de la gestión de servicios de TI, y la mayoría de los contratos revisados en la literatura las adopta sin traducirlas al dominio de la seguridad. La taxonomía de cuatro tipos de la guía NIST SP 800-55v1 permite ver por qué eso falla. Los acuerdos de nivel de servicio miden casi exclusivamente eficiencia, cuando lo que la organización necesita conocer es eficacia, esto es, si los controles están funcionando.

La segunda es que existe un vacío normativo documentable y no menor. El estándar internacional de referencia en medición de seguridad de la información no aborda, en ninguno de sus dos volúmenes, cómo medir a un proveedor tercerizado, y remite ese dominio a una guía de gestión de riesgo de cadena de suministro que a su vez no proporciona fórmulas. La única fuente pública que ofrece fórmulas explícitas control por control es la especificación de evaluación del Center for Internet Security, que no fue concebida para relaciones contractuales. El modelo propuesto opera precisamente en ese espacio intermedio.

La tercera conclusión es la de mayores consecuencias prácticas. Como el resultado del esfuerzo preventivo en seguridad es estructuralmente inobservable, todo contrato mide sustitutos, y todo sustituto se degrada al convertirse en objetivo con consecuencia económica. El caso extremo de esa degradación contradice la práctica contractual dominante en el mercado. Un acuerdo de nivel de servicio basado únicamente en penalidades por incidente incentiva al proveedor a no detectar, porque descubrir una brecha

activa investigaciones que pueden asignarle responsabilidad. El resultado está formalmente demostrado en la literatura económica de seguridad y encuentra correlato empírico en el comportamiento observado de los analistas de proveedores gestionados. Nada de esto es una objeción a medir. Es una razón para medir mejor, y la respuesta tiene tres componentes que deben operar juntos: indicadores de segundo orden que detecten la manipulación de los de primer orden, verificación independiente del reporte y una estructura contractual que combine penalidad por brecha con recompensa por detección.

La cuarta se refiere al contexto colombiano. Para las entidades vigiladas por la Superintendencia Financiera, contar con indicadores de eficacia y eficiencia no es una recomendación de buenas prácticas sino un requerimiento expreso de la Circular Básica Jurídica, acompañado de la obligación de incluir medidas de seguridad en los contratos con terceros críticos y de verificar periódicamente su cumplimiento. Existe, por tanto, un mandato regulatorio sectorial, vinculante para las entidades vigiladas y aplicable como referente de buena práctica para las demás, que este tipo de modelos ayuda a satisfacer, junto a una ausencia documentada de literatura académica en español sobre la intersección entre tercerización y medición de ciberseguridad.

La quinta la aporta el caso de aplicación, y es la que mejor resume para qué sirve el modelo: un tablero construido alrededor de un solo indicador con consecuencia económica no es una versión austera de un buen sistema de medición, sino un sistema que produce activamente una imagen equivocada. La contraparte de segundo orden opera como condición de validez de la información que el tablero presenta.

Como limitación debe declararse que se trata de una revisión documental y de una propuesta de modelo, no de una validación empírica. El caso del apartado 9 usa datos simulados y no debe confundirse con evidencia de campo. Además, el texto normativo interno de varias normas ISO citadas no pudo consultarse por tratarse de publicaciones de pago, de modo que esas fuentes se referencian por título, año y alcance oficial. La línea de trabajo futura más natural consiste en aplicar el tablero durante varios periodos en una organización real y contrastar si los indicadores de segundo orden anticipan la degradación antes que los de primer orden.

Cierro con la reflexión que ordena todo el trabajo. Medir a un proveedor de ciberseguridad no es un ejercicio de control administrativo, sino el mecanismo con el que una organización compensa la asimetría de información que ella misma creó al tercerizar, y la calidad de esa compensación depende menos de la cantidad de indicadores del tablero que del rigor con que se declaren sus límites.

Referencias

Alahmadi, B. A., Axon, L., y Martinovic, I. (2022). 99% false positives: A qualitative study of SOC analysts' perspectives on security alarms. En *Proceedings of the 31st USENIX Security Symposium* (pp. 2783–2800). USENIX Association. https://www.usenix.org/system/files/sec22summer_alahmadi.pdf

American Institute of Certified Public Accountants. (2022a). *2017 trust services criteria for security, availability, processing integrity, confidentiality, and privacy (with revised points of focus — 2022)* (TSP Section 100). https://www.aicpa-cima.com/resources/download/2017-trust-services-criteria-with-revised-points-of-focus-2022

American Institute of Certified Public Accountants. (2022b). *2018 description criteria for a description of a service organization's system in a SOC 2 report (with revised implementation guidance — 2022)* (DC Section 200). https://www.aicpa-cima.com/resources/download/get-description-criteria-for-your-organizations-soc-2-r-report

Boyens, J., Smith, A., Bartol, N., Winkler, K., Holbrook, A., y Fallon, M. (2024). *Cybersecurity supply chain risk management practices for systems and organizations* (NIST Special Publication 800-161, Rev. 1, Actualización 1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-161r1-upd1

Center for Internet Security. (2024). *CIS Controls assessment specification for Controls v8.1*. https://cas.docs.cisecurity.org/en/latest/

Cezar, A., Cavusoglu, H., y Raghunathan, S. (2014). Outsourcing information security: Contracting issues and security implications. *Management Science, 60*(3), 638–657. https://doi.org/10.1287/mnsc.2013.1763

Chamkar, S. A., Maleh, Y., y Gherabi, N. (2024). Security operations centers: Use case best practices, coverage, and gap analysis based on MITRE adversarial tactics, techniques, and common knowledge. *Journal of Cybersecurity and Privacy, 4*(4), 777–793. https://doi.org/10.3390/jcp4040036

Chew, E., Swanson, M., Stine, K., Bartol, N., Brown, A., y Robinson, W. (2008). *Performance measurement guide for information security* (NIST Special Publication 800-55, Rev. 1). National Institute of Standards and Technology. (Retirada el 4 de diciembre de 2024). https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-55r1.pdf

Cichonski, P., Millar, T., Grance, T., y Scarfone, K. (2012). *Computer security incident handling guide* (NIST Special Publication 800-61, Rev. 2). National Institute of Standards and Technology. (Retirada el 3 de abril de 2025). https://doi.org/10.6028/NIST.SP.800-61r2

Congreso de la República de Colombia. (2009, 5 de enero). *Ley 1273 de 2009, por medio de la cual se modifica el Código Penal y se crea un nuevo bien jurídico tutelado denominado "de la protección de la información y de los datos"*. Diario Oficial No. 47.223. https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=34492

Congreso de la República de Colombia. (2012, 17 de octubre). *Ley Estatutaria 1581 de 2012, por la cual se dictan disposiciones generales para la protección de datos personales*. Diario Oficial No. 48.587. https://www.funcionpublica.gov.co/eva/gestornormativo/norma.php?i=49981

Crowley, C. (2025). *SANS SOC survey 2025*. SANS Institute. https://www.sans.org/white-papers/sans-2025-soc-survey

Departamento Nacional de Planeación. (2020). *Documento CONPES 3995: Política Nacional de Confianza y Seguridad Digital*. https://normograma.mintic.gov.co/mintic/compilacion/docs/conpes_dnp_3995_2020.htm

European Union Agency for Cybersecurity. (2012). *Introduction to return on security investment: Helping CERTs assessing the cost of (lack of) security*. https://www.enisa.europa.eu/publications/introduction-to-return-on-security-investment

European Union Agency for Cybersecurity. (2021). *ENISA threat landscape for supply chain attacks*. Publications Office of the European Union. https://doi.org/10.2824/168593

European Union Agency for Cybersecurity. (2023). *Good practices for supply chain cybersecurity*. https://www.enisa.europa.eu/publications/good-practices-for-supply-chain-cybersecurity

European Union Agency for Cybersecurity. (2025). *ENISA threat landscape 2025*. Publications Office of the European Union. https://www.enisa.europa.eu/publications/enisa-threat-landscape-2025

Forum of Incident Response and Security Teams. (2019). *CSIRT services framework* (Versión 2.1). https://www.first.org/standards/frameworks/csirts/csirt_services_framework_v2.1

Forum of Incident Response and Security Teams. (s. f.). *Team types within the context of services frameworks* (Versión 1.0). https://www.first.org/standards/frameworks/csirts/team-type_1-0

Gartner. (2026, 5 de febrero). *Gartner identifies the top cybersecurity trends for 2026* [Comunicado de prensa]. https://www.gartner.com/en/newsroom/press-releases/2026-02-05-gartner-identifies-the-top-cybersecurity-trends-for-2026

Gordon, L. A., y Loeb, M. P. (2002). The economics of information security investment. *ACM Transactions on Information and System Security, 5*(4), 438–457. https://doi.org/10.1145/581271.581274

Hood, C., y Piotrowska, B. (2021). Goodhart's law and the gaming of UK public spending numbers. *Public Performance & Management Review, 44*(2), 264–280. https://doi.org/10.1080/15309576.2020.1749092

IBM Security y Ponemon Institute. (2026). *Cost of a data breach report 2026: The AI tipping point*. IBM Corporation. https://www.ibm.com/reports/data-breach

International Auditing and Assurance Standards Board. (2011). *International Standard on Assurance Engagements (ISAE) 3402: Assurance reports on controls at a*

service organization. IFAC. https://www.iaasb.org/publications/staff-overview-international-standard-assurance-engagements-isae-3402-assurance-reports-controls

International Organization for Standardization e International Electrotechnical Commission. (2018). *Information technology — Service management — Part 1: Service management system requirements* (ISO/IEC 20000-1:2018). https://www.iso.org/standard/70636.html

International Organization for Standardization e International Electrotechnical Commission. (2021). *Cybersecurity — Supplier relationships — Part 1: Overview and concepts* (ISO/IEC 27036-1:2021). https://www.iso.org/standard/82905.html

International Organization for Standardization e International Electrotechnical Commission. (2022a). *Cybersecurity — Supplier relationships — Part 2: Requirements* (ISO/IEC 27036-2:2022). https://www.iso.org/standard/82060.html

International Organization for Standardization e International Electrotechnical Commission. (2022b). *Information security, cybersecurity and privacy protection — Information security controls* (ISO/IEC 27002:2022). https://www.iso.org/standard/75652.html

International Organization for Standardization e International Electrotechnical Commission. (2022c). *Information security, cybersecurity and privacy protection — Information security management systems — Requirements* (ISO/IEC 27001:2022). https://www.iso.org/standard/27001

International Organization for Standardization e International Electrotechnical Commission. (2024). *Information security, cybersecurity and privacy protection — Requirements for bodies providing audit and certification of information security management systems — Part 1: General* (ISO/IEC 27006-1:2024). https://www.iso.org/standard/82908.html

Mandiant. (2026). *M-Trends 2026: Data, insights, and strategies from the frontlines*. Google Cloud. https://cloud.google.com/blog/topics/threat-intelligence/m-trends-2026

Marquez, J. (2021, 5 de enero). Indicator management: A question that should start with why. *ISACA Industry News*. https://www.isaca.org/resources/news-and-trends/industry-news/2021/indicator-management-a-question-that-should-start-with-why

Ministerio de Tecnologías de la Información y las Comunicaciones. (2021). *Guía para la gestión y clasificación de incidentes de seguridad de la información* (Guía 21). https://gobiernodigital.mintic.gov.co/692/articles-5482_G21_Gestion_Incidentes.pdf

National Institute of Standards and Technology. (2024). *The NIST Cybersecurity Framework (CSF) 2.0* (NIST CSWP 29). U.S. Department of Commerce. https://doi.org/10.6028/NIST.CSWP.29

Nelson, A., Rekhi, S., Souppaya, M., y Scarfone, K. (2025). *Incident response recommendations and considerations for cybersecurity risk management: A CSF*

2.0 community profile (NIST Special Publication 800-61, Rev. 3). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-61r3

Nicolazzo, S., Nocera, A., y Pedrycz, W. (2024). *Service level agreements and security SLA: A comprehensive survey* [Preprint]. arXiv. https://arxiv.org/abs/2405.00009

Rumburg, J., y Zbikowski, E. (2012). *The seven most important performance indicators for the service desk*. HDI / MetricNet. https://www.thinkhdi.com/~/media/HDICorp/Files/Library-Archive/Rumburg_SevenKPIs.pdf

Sanguino Reyes, M. R., y Cuesta Quintero, B. (2023). Tercerización de servicios de tecnologías de información: modelo de gobierno para las universidades públicas. *Revista CEA, 9*(21), e2398. https://doi.org/10.22430/24223182.2398

Schroeder, K., Trinh, H., y Pillitteri, V. (2024a). *Measurement guide for information security: Volume 1 — Identifying and selecting measures* (NIST Special Publication 800-55v1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-55v1

Schroeder, K., Trinh, H., y Pillitteri, V. (2024b). *Measurement guide for information security: Volume 2 — Developing an information security measurement program* (NIST Special Publication 800-55v2). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-55v2

Superintendencia Financiera de Colombia. (2020). *Circular Básica Jurídica, Parte I, Título IV, Capítulo V: Requerimientos mínimos para la gestión de la seguridad de la información y la ciberseguridad* (modificado por la Circular Externa 033 de

2020). https://www.superfinanciera.gov.co/publicaciones/10083444/normativanormativa-generalcircular-basica-juridica-ce-parte-i-instrucciones-generales-aplicables-a-las-entidades-vigiladas-10083444/

Swanson, M., Bowen, P., Phillips, A. W., Gallup, D., y Lynes, D. (2010). *Contingency planning guide for federal information systems* (NIST Special Publication 800-34, Rev. 1). National Institute of Standards and Technology. https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-34r1.pdf

Tammineedi, R. L. S. (2018). Integrating KRIs and KPIs for effective technology risk management. *ISACA Journal, 4*. https://www.isaca.org/resources/isaca-journal/issues/2018/volume-4/integrating-kris-and-kpis-for-effective-technology-risk-management

Van Os, R. (2026). *SOC maturity report: Trends and insights in global SOC maturity* (2.ª ed. anual). SOC-CMM. https://www.soc-cmm.com/img/upload/files/66-soc-maturity-report-2026.pdf

Verizon. (2025). *2025 data breach investigations report*. Verizon Business. https://www.verizon.com/business/resources/reports/2025-dbir-data-breach-investigations-report.pdf

Verizon. (2026). *2026 data breach investigations report*. Verizon Business. https://www.verizon.com/business/resources/reports/2026-dbir-data-breach-investigations-report.pdf